

Young scientists deserve better of the system (yet again)

A high-level report on careers in the life sciences raises often-repeated worries about employment practices afflicting graduate students and postdoctoral researchers in the United States. A major rethink is required.

This week sees more fine words added to the mountain of pronouncements on the vexed issue of careers of young scientists, and in particular the consequences of the fact that there is a glut of talented young researchers with little prospect of ever obtaining tenured positions. But who among those potentially accountable for changing things for the better will lift a finger to improve matters?

Every scientific discipline suffers from the problem to some extent. The latest report (see page 103), from the US National Research Council (NRC), examines the situation of young life scientists in the United States. Even there, where private and public investment in research is climbing ever upward above the 1997 total of almost \$40 billion, advertised permanent positions are hugely over-subscribed. Meanwhile, researchers in contract positions get closer and closer towards senility before finding permanent scientific posts — or having to leave research without accumulating reasonable employment benefits.

A sobering statistic in the report reveals that, whereas 61 per cent of those who obtained a PhD in the early 1960s achieved tenured university positions within ten years, that figure has fallen to 38 per cent for those qualifying in the mid-1980s. One can only contemplate with gloom the possible statistic ten years hence.

Four years ago, a report from the National Academy of Sciences drew attention to similar statistics across all scientific disciplines, but took a much more relaxed view of the situation. It concluded, for example, that employment outside academic institutions does not necessarily constitute a misuse of doctoral programmes. Since then, funding for the life sciences has increased at a rate that may have exacerbated the plight of young scientists.

Dubious alternatives

Many of the NRC report's messages have been heard often before: problems of dissatisfaction from over-optimistic career expectations, the importance of the master's degree as a useful alternative to the doctorate, the high proportion of immigrant researchers in the United States, and the need for better information for graduates on the difficulties of postdoctoral employment.

But the report makes important new headway in at least one respect: it introduces some justifiable scepticism towards the idea that there are easy career alternatives outside science for those with PhDs or postdoctoral experience. A wider question becomes even more pressing than hitherto: should so much in the way of research funds be used to provide training that is of dubious relevance to most ultimate career destinations?

The NRC report is also welcome because of the attention it brings to the tasks that young scientists are given within US laboratories which in other countries, notably Germany, would be carried out by permanent technical staff. But it is not clear why institutions currently enjoying a surplus of cheap labour should adopt an expensive increase in their permanent overheads. Their research can thrive even though young scientists are being diverted from research goals by tasks that do not directly benefit their careers. After all, publications, rather than experi-

ence in teaching or tending and developing sophisticated equipment, are what counts when trying to move outwards or upwards.

The report advocates a halt in the expansion of PhD programmes. However, as is also the case with many of the report's other suggestions, there is no central authority in the United States that can directly enforce such a policy. And there are good reasons why the upward pressure will persist. In the US system, small universities with big aspirations are not going to be prohibited from pursuing their ambitions, and congressional politics will continue to encourage a wide distribution of centres of research excellence.

Changing thresholds

Underneath the surface of the career problems of the postdoc sits the highly charged issue of immigration. The report recommends against direct restrictions on immigration. It may be that it has erred towards timidity in avoiding that option, and that a deeper investigation of numbers and the possible consequences of tightened quotas would be worthwhile. Such solutions may be suspect for reasons of practice, politics and principle, but that will not stop some people promoting tighter restrictions, and it is surely advisable to have a full analysis on the table.

Another approach could be to raise the threshold of acceptance into postgraduate research — an obvious though slow way forward in meeting the report's recommendation of an immediate freezing of numbers, but one that at least does not discriminate against foreign applicants. Why not insist on higher levels of qualification or demonstrated ability, perhaps including a master's degree for some categories, as a condition of funding young scientists for research?

Beneath these aspects of the problem lies a tension, particularly characteristic of the United States, between local pressures and institutional autonomy on the one hand, and limited central power on the other. Who is going to encourage institutions to change their employment practices? There seem to be two possible sources of leverage. One, highly controversial and in principle, at least, wholly undesirable, would be restrictions on immigration. The other lies at the door of the National Institutes of Health (NIH). The rate of NIH funding increases in recent years is fuelling the boom in applicants and temporary employment which the NRC is highlighting in its report. Four years ago, the National Academy of Sciences concluded that the problem did not require significant changes in practice. That is no longer tenable. The NIH holds no direct regulatory power, but can nonetheless exercise enormous influence over the laboratories it supports. By introducing consideration of employment structures and practices into its funding criteria, it could exert pressure for overdue change in laboratories across the United States.

Such a policy would represent a significant and contentious shift. But the problems facing young scientists are both chronic and acute. A fresh approach is required to prevent the current boom in the life sciences from degenerating further into a boulevard of broken dreams for too many young investigators. □



Jobs crisis sparks call for freeze in number of PhD students in US

[WASHINGTON] The number of life sciences PhD students in the United States should be frozen at current levels, recommends a report by the National Research Council. *Trends in the Early Careers of Life Scientists*, released today (10 September), says that increasing numbers of graduate students and postdoctoral fellows are failing to find long-term jobs.

Nearly four times more PhDs are being produced each year than in 1963, with a 42 per cent rise between 1987 and 1996. Almost all the growth is in biomedical fields, and much is accounted for by foreign scientists. But jobs in the academic community, government and industry have not kept pace. There have been increases in the length of time taken to complete a PhD, the average age at graduation, the numbers of young scientists taking up postdoctoral fellowships and the length of time spent as postdocs.

In 1973, about 11 per cent of PhDs still held postdoctoral or other non-faculty jobs at universities five to six years after graduating, or were outside science altogether. By 1995, that number had risen to 38 per cent.

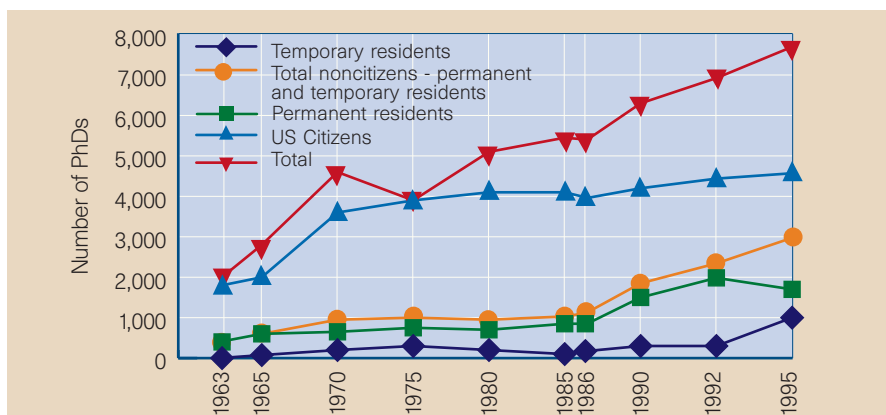
The report says that this situation has produced a "crisis in expectation" among young scientists. "The feelings of disappointment, frustration and even despair are palpable in the laboratories of academic centers," it says. "Further increase in the competition could discourage the best from entering the field."

The authors argue that even projected increases in government funding — for instance, an anticipated doubling of the budget of the National Institutes of Health — are unlikely to ensure any mitigating growth in permanent jobs. Nor, they say, can industry be counted on to hire more people, given recent trends.

They recommend that life sciences PhD programmes should not be expanded, and that no new programmes be started except in "rare and special circumstances", such as encouraging the education of under-represented minorities.

"The current rate of growth can no longer be justified, and the premises that have produced it must be re-examined," the report says. "The system is delaying independence and muffling creativity at perhaps the most productive phase of the scientist's life."

If growth is not stopped, the report notes, the situation will only worsen: for instance, if the 5.1 per cent growth rate in PhD graduates between 1995 and 1996 were continued, the number of PhD graduates would double in 14 years.



Population explosion: foreign students account for much of the growth in US life sciences PhDs.

William Brinkley, the president of the Federation of American Societies for Experimental Biology, and a member of the panel that wrote the report, says: "The hard facts are that across the life sciences we have a huge reservoir of outstanding people, capable of filling the needs that we see in the growth of science in the next few years."

The report urges universities to move towards using permanent, non-PhD technicians to carry out work now being done by graduate students and postdoctoral fellows. It also suggests that they consider restricting the numbers of graduate students supported by research grants.

Shirley Tilghman, a professor of molecular biology and a Howard Hughes Medical Institute investigator at Princeton University, who chaired the 16-member panel that wrote the report, blames the structure of the US research enterprise, in which senior scientists rely on cadres of young graduate students and some 20,000 postdocs to staff their laboratories at bargain rates.

"We are creating this working class who are absolutely essential to the conduct of research, and yet they are spending many years in positions which we would not consider on some kind of a career trajectory."

Tilghman likens the situation of postdocs to jets above an airport: "They are circling, using up fuel and waiting for their turn to land. We have to think hard about the ways in which we have allowed what were supposed to be training positions to become these kinds of holding-pattern positions."

But some in the field question whether a moratorium on PhD programme growth would be accepted by institutions and senior scientists now heavily reliant on young scientists for carrying out research.

"It's difficult to tell people to stop training people when they look at a trainee and think:

'this is how we get our work done,'" says Frank Solomon, a professor of biology at the Massachusetts Institute of Technology who chairs the education committee of the American Society for Cell Biology.

Others question whether the crisis described in the report has reached the proportions that would require a freeze on PhD programmes. "We cannot predict future job opportunities and therefore we cannot say when we're going to need to enter the steady state. We don't want to turn away bright young people because of a negative press," says Susan Gerbi, chair of the department of molecular biology, cell biology and biochemistry at Brown University in Providence, Rhode Island.

The report says that graduate programmes should be required to confront prospective students with the hard facts of their career prospects, by providing data on the careers of PhD graduates over the previous 10 years. The information "will have the salutary effect of letting market forces control the rate of entry into the profession."

And the report urges government agencies to shift towards funding more broad-based training grants and individual fellowships, with fewer of the research grants that have almost entirely fuelled the growth in the number of PhD trainees over the past 10 years. While attractive to institutions as sources of funded students, these do less for the student, the report says.

The report dismisses the notion that PhD programmes should be broadened to prepare graduates for alternative careers that, it maintains, do not require a PhD.

The report says foreigners accounted for most of the increase in PhD graduates since 1987. But it urges that this should not lead to "arbitrary limitations on the number of visas issued for foreign students". **Meredith Wadman**

Retraining scheme could go Europe-wide

[MUNICH] An experimental project to retrain unemployed scientists from the former East Germany appears so successful that European Commission officials are now "looking favourably" at backing similar schemes elsewhere in Europe.

The scheme, using European Union (EU) funds, was introduced last year by the Max Delbrück Centre for Molecular Medicine (MDC) in Berlin, where 60 scientists are employed using money from the European Social Fund (ESF), part of the EU's subsidies for poor regions (see *Nature* 388, 109; 1997).

After only one year of the three-year retraining scheme, more than one in ten have found permanent positions or won independent grant money. A further DM10 million (US\$6 million) was approved earlier this year to retrain 46 researchers at Berlin's Adlershof Science Park. Other institutes in Berlin, Brandenburg and Sachsen-Anhalt have also applied for ESF money.

EU representatives told a meeting in Berlin that the encouraging results mean the scheme may be expanded to other regions after 2000. Cuts in public spending mean that many countries, such as Spain and Portugal, face major difficulties in providing full-time research posts for researchers who have completed their postdoctoral training.

ESF training programmes have traditionally been restricted to unqualified work-

ers. But Karl Freese, of the Association of Biomedical Research, a German lobby group representing 15 non-university biomedical research institutes, is one of many campaigning for qualified workers to be included.

Freese argues that scientists educated in the former East Germany, for example, did not have the chance to learn modern scientific techniques. He is worried that several thousand face dismissal when their non-renewable short-term contracts expire in the autumn. "Unemployment among scientists is a waste of human creativity," says Freese.

The commission appears to have been won round. But Hans Keirat, an EU official who oversees allocation of social funds for east Germany, warns that before other countries can take advantage of it, the commission "wants to be convinced" that the Berlin programmes have really reduced long-term unemployment. Marion Bimmler, of the MDC's personnel department, accepts that the centre needs more time to find out if this goal can be achieved. "But we are very optimistic about it," she says, denying that the scheme is merely a "welfare act".

Participants continue to work in established research groups while retraining in new techniques. According to Bimmler, cooperation between the 'trainees' and staff researchers has been excellent. "ESF-funded scientists have contributed to many publica-

tions and patents and have even co-founded new companies," she says.

Karla Köpke, for example, a 47-year-old mathematician, worked for the Academy of Sciences before reunification and then at the Humboldt University in Berlin, until her funding ran out in 1996. Last year she joined a research group at the MDC analysing genes underlying drug addiction. She developed a mathematical method for analysing the polymorphic profiles of candidate genes, which was not possible by conventional methods. "It was because of Köpke's work that we managed to develop a classification of various genetic profiles," says Margarete Hoehe, head of the research group. Köpke's method is to be patented.

If the Berlin results continue to be positive, the chances of creating a general scheme for retraining unemployed scientists are high. Changes in the distribution of EU subsidies to poor regions, which absorb a third of the EU's total budget, are being negotiated. After 2000 they are likely to be used to address general issues such as unemployment throughout the EU.

Biochemist Christof Tannert, a socialist member of the European Parliament from east Germany, says: "Social funds alone cannot solve unemployment among scientists. But the fact that their doors are now open to scientists will help."

Quirin Schiermeier

Stock market crash threatens start-up biotech companies

[SAN FRANCISCO] The plunge in the US stock market may finally knock biotechnology shares out of the doldrums and trigger more investor interest, experts suggest.

At the same time, they warn that the industry may suffer enough in the short term to propel the weakest, particularly new companies, into mergers or out of business.

In the two weeks up to 4 September US investors suffered their worst episode since the ten days before the stock market crashed in 1987. The Dow Jones Industrial Average plunged 512 points on 31 August, and by the end of the week was down 3.4 per cent for the year. The Nasdaq index, which more closely reflects the performance of technology and biotechnology stocks, dropped 73.16 points over the week, 0.2 per cent down for the year. Another measure, the Russell index of small company stocks, rose slightly, but is still down 20.6 per cent for the year.

Biotechnology stocks have been weak since last October, when concerns about the crisis in Asia hit investors' nerves, but stock market analysts saw reason for optimism.

The strong performance of the Dow Industrial Average and the S&P 500 had taken away the incentive to hold smaller

stocks, says Dave Stone, managing director and biotechnology analyst at S. G. Cowen Securities Corp. in Boston. Now investors may be drawn to biotechnology by the industry's improving fundamentals.

"Shaking investors off the bandwagon is probably not a bad thing for the biotech sector," says Stone.

In fact, a few biotech companies bounced back up later last week. These included industry leader Chiron, based in Emeryville, California; Vertex Pharmaceuticals, a specialist in structure-based drug design for viral diseases, in Cambridge, Massachusetts; and AgriBiotech, a Las Vegas forage and turf seed company. Shares in the San Diego-based agricultural biotechnology company Mycogen rose \$5.72 to \$27.66 a share. On 1 September, Dow Chemical said it would buy the 32 per cent of Mycogen it did not own for \$28 a share, or nearly \$325 million.

James McCamant, editor of the *Medical Technology Stock Letter*, has urged readers to load up immediately on biotech stocks. Biotechnology companies are barely affected by the worldwide economic swings because they serve the burgeoning US medical sector, he wrote. Wall Street has largely

ignored strong sales of new products such as Agouron's AIDS drug Viracept, good clinical trial results, broad-based development programmes and smart corporate deals — creating some compelling bargains.

Those bargain prices have made it hard for companies to raise capital for operations, however. Misha Petkevich, managing partner in Petkevich and Partners in San Francisco, predicted increasing difficulty for biotechnology companies not seen as solid in technology and management. Those with a broad technology platform, a product close to market, or involved in a service function such as gene identification may benefit from a period of consolidated investor interest while others fall by the wayside, he said.

Petkevich said the outlook for new biotechnology companies was less encouraging. While venture capitalists have plenty of cash right now, "there's a significant exodus from being involved in start-up biotech companies," he said. Venture capitalists have been lukewarm towards bioscience for the past year, and the recent stock market volatility has not helped to increase their interest.

Sally Lehrman

Neutrino study 'needs transatlantic effort'

[WASHINGTON] US particle physicists are trying to persuade their European rivals to join them in a joint effort to study neutrinos.

Experiments involving neutrinos are extraordinarily difficult, say the physicists who spend their lives scrutinizing these virtually undetectable particles. It therefore makes sense for rival teams to adopt different experimental approaches. But US neutrino scientists are left wondering why cash-strapped high-energy physics programmes on both sides of the Atlantic are pursuing separate, but highly similar, projects to measure the mass of a neutrino.

Both proposed projects will harness protons from accelerators to generate beams of muon neutrinos, which they will send underground, at an angle of inclination of around 5 degrees, towards massive detectors 750 km away. This will establish how many of the muon neutrinos 'oscillate' (into tau neutrinos) during the journey.

One project would send the neutrinos from Fermilab in Illinois to northern Minnesota, while the other would start its beam at CERN, the European Laboratory for Particle Physics in Geneva, and observe the result at the Gran Sasso Laboratory near Rome.

Oscillation between any two of the three classes of neutrino — electron, muon and tau — would indicate that they have mass. This would require the modification of the prevalent theory of particle physics, the Standard Model, which holds that the neutrino is massless.

Oscillation of neutrinos generated naturally in the atmosphere, as they pass through the Earth, was recently observed at the SuperKamiokande experiment in Japan (see *Nature* **391**, 123; 1998). An experiment to generate neutrinos at KEK, the Japanese particle accelerator, and direct them underground to the Kamiokande detector 250 km away, will begin in January. The United States and Europe are both planning similar but more powerful experiments, with more energetic neutrinos sent over a longer distance, in a bid to confirm and measure their mass.

The US project, Neutrinos at the Main Injector (NUMI), has been approved by the Department of Energy, and construction of its \$76 million neutrino beamline is commencing at Fermilab. The beam will send muon neutrons through a 'near' detector at Fermilab to a 'far' one 730 km away at Soudan, Minnesota, down a disused iron mine. The experimental side of the project, known as the Main Injector Oscillation Search (MINOS), will cost a further \$60 million.

European plans for a similar project have been around for years, and room to construct a suitably inclined neutrino beamline is built into the design of CERN's new Large



US neutrino detector should remain part of national effort, say Europeans.

Hadron Collider. "Everyone agrees it's a great idea but we're not sure where to find the money," says John Ellis, head of the theory group at CERN.

Although CERN can provide no cash, the Italian National Institute for Nuclear Physics (INFN) has committed 60 per cent of the \$50 million it would cost to build the beamline in Switzerland, and additional money for ICARUS, the detector at the INFN Gran Sasso Laboratory.

Most proposals received so far for the Gran Sasso experiment are from Italian groups, says Luigi di Lella, head of the NOMAD neutrino experiment at CERN. At a workshop in Amsterdam this week, its supporters will seek to broaden its appeal beyond Italy. "We hope other institutes will join in, and that other countries will share in the cost of the beam," says di Lella.

But US physicists will encourage Europeans to join them instead, and perhaps to use whatever money they can raise to add an experiment of their own using the NUMI beam.

Stan Wojcicki, a professor of physics at Stanford University in California and a spokesman for MINOS, says that European neutrino scientists should join the US project, which already has UK and Russian collaborators. "I think it makes perfect sense," he says. "The Europeans clearly feel they'd have more impact if an experiment was done in Europe, but there is a shortage of money everywhere."

European scientists should consider making moves to collaborate immediately, rather than "waiting in the wings" while Gran Sasso seeks full funding, he says. "Unfortunately the right time for them to join [MINOS] is now: in two or three years, things will be cast in stone."

Supporters of the European experiment are not convinced of the justification for a single international experiment. The costs are simply not high enough, says Enzo Iarocco, president of INFN. Iarocco and other supporters of the Gran Sasso plan also say that, because neutrino physics is so subtle, it is scientifically preferable to have two systems, as a check. Iarocco also points out that Gran Sasso is "a unique facility, designed 15 years ago to have its underground laboratory orientated in the direction of CERN — precisely to catch the neutrino beam it hoped would one day be sent".

The US group counters that the European proposal is too weak to compete effectively. Robert Bernstein, one of the MINOS team, says he believes in a "let a thousand flowers bloom" approach to neutrino experiments — but not in this case. He says MINOS will detect 38,000 events a year when it starts operating in 2002, whereas ICARUS will open later and detect fewer than 2,000 events. With both a near and a far detector, MINOS will be able to count muon neutrinos at both ends of the journey, as well as looking for tau neutrinos at the end. The European proposal has only a far detector, although options for a near detector will be discussed at Amsterdam.

But a CERN spokesman says the two experiments are equivalent but not directly comparable, because the larger flux of neutrinos achieved by MINOS is compensated by the higher energy of the ICARUS beam.

Other physicists at Fermilab, which is heavily involved in the \$450 million contribution to the Large Hadron Collider (see *Nature* **394**, 611; 1998), say Europe should help with NUMI in reciprocation. But agencies on both sides of the Atlantic believe that no such reciprocation is due on a project of NUMI's modest size.

Colin Macilwain

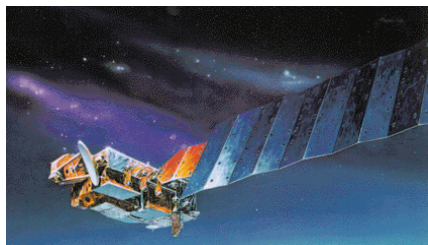
NASA charts Earth observation strategy

[WASHINGTON] About 100 scientists gathered last month to begin charting a course for the next generation of low-cost Earth observation satellites, to follow the last of NASA's Earth Observing System (EOS) spacecraft scheduled for launch in 2002.

NASA called the invitation-only workshop to begin forming a consensus among scientists about missions to be flown from 2003 to 2010. The meeting was held near Washington DC and included observers from Europe, Japan, Canada and Brazil.

A panel led by Charles Kennel will report the results of the workshop by mid-October. Kennel was formerly the head of NASA's Earth science programme and is now director of the Scripps Institution of Oceanography in San Diego.

Before the meeting NASA received about



Eyes in space: an armada of Earth-observing satellites could be heading for orbit next century.

100 responses from scientists to a request for mission concepts. These were whittled down to the 23 most promising ideas by six discipline panels for climate studies, atmospheric chemistry, hydrology, oceans and ice, land cover and terrestrial ecosystems, and geodynamics and geology.

Programme managers at the space agency produced an integrated "strawman" list of missions for consideration at the workshop. James Dodge of NASA's Earth science office says the list was only a "going-in position", meant to stimulate discussion, and is almost certain to change before the agency begins requesting detailed proposals for missions, perhaps as early as 2000.

Those attending the workshop were given certain ground rules. Even the largest of the next generation of spacecraft are expected to cost 30 per cent less than the first generation of EOS missions. International and commercial partnerships are encouraged, and any proposal should assume cooperation with the National Polar-orbiting Operational Environmental Satellite System (NPOESS), the civilian-military weather satellite system that will see its first launch in 2009.

But many scientists are sceptical that the interests of operational weather forecasters and researchers will prove compatible in a joint programme. Also to be determined is what role new technologies will play in lowering mission costs while improving performance. NASA has begun an "incubator" programme to develop advanced instruments for Earth-viewing spacecraft, with the first grants to be announced this month.

The strawman list includes ideas for large spacecraft as well as small 'Earth probe' missions, and would balance long-term, systematic measurements of the kind conducted in the EOS series with more exploratory research satellites.

Among the concepts proposed for earliest development are a Land Cover/Land Use Inventory mission for launch in 2005 as a follow-on to Landsat 7; a Climate Variability and Trend satellite in 2006 to bridge the gap between the EOS-PM mission and the launch of the first NPOESS satellite; and a 2005 mission to monitor biological productivity of the oceans and land, which would bridge the gap between the EOS-AM spacecraft and the first NPOESS satellite.

NASA managers concede that a two-day meeting was hardly enough time to form a meaningful consensus on priorities. But Mark Abbott of Oregon State University says: "This isn't the final story — you'll see it evolve over the next year." NASA intends to have the Kennel group's recommendations reviewed by its own Earth science advisory committee and by the National Academy of Sciences.

Abbott, who chairs an advisory committee for EOS scientific payloads, says those attending the workshop represented a "much broader group" than just EOS investigators, and the "science questions were more sharply focused" because knowledge has advanced since the EOS missions were planned more than a decade ago.

Tony Reichhardt

Call for Europe-wide public health agency

[PARIS] A group of international scientists meets for the first time this week, in Montpellier, France, to advocate the creation of a European equivalent of the US Centers for Disease Control (CDC). But political support for the proposal seems unlikely to be forthcoming at present, given the reluctance of the European Union's member states to cede sovereignty over public health policy to the European Commission.

The proposal for a European Centre for Control of Infectious Diseases (ECCID) is being championed by Michel Tibayrenc, from the Centre d'Etude sur le Polymorphisme des Micro-organismes in Montpellier. It has gained the support of several scientific organizations, including the European Society of Clinical Microbiology and Infectious Diseases. A committee of two dozen scientists has been established to promote ECCID, including Francisco Ayala from the University of California at Irvine; Dan Colley, head of the CDC's Division of Parasitic Diseases; and Koussay Dellagi, director general of the Pasteur Institute in Tunis.

The CDC was created in 1946 and has grown into a national disease prevention agency with nearly 7,000 staff. It runs large programmes for detecting and investigating health problems, carrying out research and control, training health professionals and promoting common policies in public health. An equivalent made up of the European Union and the countries of eastern Europe is needed in particular, argues Tibayrenc, to counter the threat of emerging and re-emerging diseases.

"In Europe, we have very poor coordination [of public health research and policy]," says Jean-Claude Piffaretti,

chairman of the Swiss Society of Microbiology and a member of the new committee. He adds that a European CDC could create a much needed "big fundamental research effort in infectious diseases" and provide a focus for European efforts in developing countries.

The need for a more coordinated European effort in public health seems to attract a broad consensus among researchers and policymakers. But political support for a centralized public health agency is missing.

"Public health is politically very sensitive; we are talking about communicable diseases and big stakes in the economy of a nation," says one official from the European Commission's public health directorate. Individual member states would be reluctant to have a supranational body scrutinizing their national health systems, he says.

Indeed, the principle of a centralized organization, which was promoted by the European Parliament, was recently rejected by the Council of Ministers. Under the compromise agreement they reached, a network for the epidemiological surveillance and control of communicable diseases will link public health centres in member states with each other and with the commission's services.

The ECCID proposal draws attention to gaps in Europe's public health systems. Although the United Kingdom has a well developed Public Health Laboratory Service, Germany and other countries are widely considered to have poor public health systems, for example in surveillance, food safety and vaccination, compared with the United States.

Declan Butler

Rapid shake-up at German health institute

[MUNICH] The Robert Koch Institute (RKI) in Berlin, which belongs to the health ministry, has undergone one of the fastest and most extensive restructuring exercises ever performed by a German research institute.

Within six months of a highly critical evaluation by the Wissenschaftsrat, Germany's science council, the RKI has flattened its hierarchical structure, slashed the number of its research departments and groups, and radically altered its role as collector and disseminator of epidemiological data.

It used consulting companies to help its 600-odd employees adapt psychologically to the changes in their working practices.

The RKI's acting director, Reinhard Kurth, says the reform should serve as a model for other ministry-owned research institutes requiring external evaluation.

The health ministry was forced to review the structure of its institutes in 1993, when information about HIV-contaminated blood failed to reach the minister, Horst Seehofer. Seehofer dissolved the ministry's public health bureau, and merged the bureau's six institutes into three — including the RKI — which are now responsible to the ministry.

The Wissenschaftsrat was asked to evaluate the three institutes. Its report, published late last year, said the RKI was performing poorly in terms of both research and other functions, including assembling epidemiological information for the ministry. It said the institute should close if it could not be radically reformed.

Kurth, who was appointed acting direc-



tor in 1996, acted rapidly to address these criticisms. RKI research will now concentrate on three areas: antibiotic resistance, Borna virus infection, and Creutzfeldt-Jakob disease. The number of research groups has been cut from 53 to 10.

The RKI will also be responsible for collecting, interpreting and disseminating epidemiological data on both non-infectious and infectious diseases. It will also improve the notoriously poor quality of epidemiological data in federal Germany, where 440 authorities collect local data, and even data on diseases like measles and rubella are unreliable. The RKI will be helped by a law under discussion in parliament that will impose sanctions on doctors who do not report cases of infectious diseases.

As a result of these changes, nearly all the RKI's staff, who are employed on permanent

civil-service contracts, will have to change the orientation of their jobs. They must also learn to work in a non-hierarchical environment, which the Wissenschaftsrat had recommended to increase flexibility and improve efficiency.

Kurth hired external consulting companies to run workshops to help with both the psychological and the practical aspects of the changes. The exercise has been expensive, he says, but the DM500,000 (US\$290,000) so far spent on workshops was "money well spent".

The RKI practice of employing staff on life-long contracts means that most staff are relatively old. This makes accepting changes much more difficult, says Kurth, "but it also means that we desperately need new blood to make sure the institute has a strong future".

He has persuaded the finance ministry to pay for two young scientists' groups to be set up for five years. The ministry says it would consider adding more in the 2000 budget.

Bärbel-Maria Bellach, an east German mathematician who is director of the newly styled department of epidemiology at the RKI, says the changes were hard for the staff to cope with, but that the workshops "really helped a lot".

She says the worst time was the three months of waiting to hear exactly what the new structure would be.

An atmosphere of suspicion remains over the distribution of jobs, admits Bellach, but younger scientists are optimistic that the institute has a good chance of becoming internationally competitive. **Alison Abbott**

Japanese science ministry expects large funding boost

[TOKYO] Japan's Ministry of Education, Science, Sports and Culture (Monbusho) is expecting a substantial increase in its science-related budget under the government's special one-off appropriation aimed at developing infrastructure for telecommunications, science, technology and the environment for the next century.

This is on top of its modest 1999 budget request for an increase of 1.8 per cent submitted last week, and will support university infrastructure and postdoctoral programmes (see *Nature* 395, 3; 1998).

Monbusho's core budget will receive

¥28.2 billion (US\$213 million) from the appropriation. The money will top up its programme for upgrading the capacity of information networks linking universities across Japan and improving database and electronic libraries at universities.

It will also boost Monbusho's 'research for the future' programme of large grants that include support for postdoctoral research assistants. The funds would increase by nearly 23 per cent, and the number of five-year industrially related projects would increase by a fifth to around 300. Each project presently receives between

¥50 million and ¥300 million a year.

Monbusho also plans to expand next year's postdoctoral fellowship programmes in line with the government's scheme to increase the number of postdocs in Japan to 10,000 by the end the decade. Next year the number of fellowships will increase by nearly a quarter to 1,600.

The 1999 budget requests reduced funding for space science and high-energy accelerators by 13.6 and 13.9 per cent, respectively. But Monbusho has asked for a quantum leap in budget from ¥5 million to ¥10.4 billion for an ambitious project to detect neutrinos. This will be a collaborative effort between the National Laboratory for High Energy Physics (KEK) and the Institute of Cosmic Ray Research at Tokyo University.

The Long Baseline Neutrino Oscillation Experiment will send neutrinos from KEK's 12-GeV proton synchrotron accelerator through a 250-kilometre underground tunnel to detectors at the SuperKamiokande observatory in central Japan. **Asako Saegusa**

Highlights of Japan's Ministry of Education, Science, Sports and Culture budget request for fiscal year 1999 (in ¥ billion)

	Request	% increase over 1998
Grants in aid for scientific research	137	16.2
Centre of excellence programme	14.4	19
Research for the future programme	26.8	22.9
Postdoctoral fellowships	17.9	11.3
Joint research with industry	118.2	11.5
Neutrinos	10.4	207,500

US unlikely to swallow plan for a food safety supremo

[WASHINGTON] The three heads of a new presidential council appear unlikely to act on a call for a single authority to be put in charge of the US government's food safety efforts.

The President's Council on Food Safety must respond to the report, by the Institute of Medicine and the National Research Council, within six months. At a recent press conference the joint chairs — Donna Shalala, the Secretary of Health and Human Services; Dan Glickman, the Secretary of Agriculture; and Neal Lane, the Assistant to the President for Science and Technology — pledged to consider its recommendations seriously.

But their comments made clear that they are unlikely to embrace its proposal for a single, powerful official, possibly heading a unified agency. The report recommended that a presidentially appointed official should control the country's food safety budget and policy, to unify and render science-based what it called a fragmented system "in critical need of attention" (see *Nature* 394, 822; 1998).

The three chairs say that existing policies need to be coordinated. "What we're interested in is ratcheting up and improving the quality of food safety ... not automatically [concluding] that what you need is a powerful new bureaucracy," said Shalala.

Lane said that too many different fields of science were involved. "You can't expect all that to appear in one agency. First, it would be huge, and second, you would have pulled all the pieces out of the other agencies that

need these ... scientific and technological underpinnings for their own missions."

Within a week of the report's release President Clinton responded by establishing the Council on Food Safety, charging it with coordinating the efforts of a dozen US agencies involved in ensuring food safety.

In addition to its three co-chairs, the council includes five cabinet-level and White House officials. It will be responsible for developing a "comprehensive strategic plan" for national food safety — taking into account the report's recommendations — and will present a unified budget to Congress. But it will not have control over food safety funds; this remains with the agencies.

Opposition to the single-agency idea goes beyond the council. Political observers widely agree that the Republican Congress would be highly unlikely to enact a law establishing a food safety bureaucracy or leader. The idea is also opposed by the food industry.

"The President can't really do what the [National Research Council] has recommended without help from Congress, so the Clinton administration is going as far as it can" by establishing the council, says Caroline Smith DeWaal, director of food safety at the Center for Science in the Public Interest.

Timothy Willard, of the National Food Processors Association, says the new council is a step forward. "We're supportive as much of what wasn't done as what was done in creating this," he says.

Meredith Wadman

Row in India over rules on animal experiments

[NEW DELHI] The National Academy of Sciences of India has urged the prime minister to intervene to prevent the implementation of new rules on using animals in research.

The controversial rules, which scientists say would strangle research with red tape, were framed by a committee led by Maneka Gandhi, minister for welfare and an ardent animal activist (see *Nature* 394, 516; 1998).

"Some of the rules formulated by this committee are likely to stall essential animal-based research rather than regulating it rationally," Prakash Tendon, the academy's president, told the prime minister.

Heads of India's biomedical research agencies and secretaries of scientific departments who met last week agreed unanimously that the legislation should be reversed. In a letter to Gandhi, Nirmal Ganguli, chief of the Indian Council of Medical Research (ICMR), warned of "far-reaching consequences" if the rules were implemented, and strongly urged her to wait until they had been debated by the scientific community.

Under the new regulations, no research institute would be able to acquire animals without the committee's permission, or begin any experiment without its clearance. Funding agencies like ICMR would have to submit detailed monthly reports of the experiments they are funding and the number of animals used. Researchers would not be allowed to import animals.

"If the guidelines are accepted by the government, much biological research in India will come to a standstill as scientists grapple with a mountain of red tape," said the Indian journal *Current Science* in an editorial. Gandhi, however, argues that scientists are killing animals unnecessarily, especially in research by pharmaceutical companies.

Vulimiri Ramalingaswami, former president of ICMR and now professor at the All-India Institute of Medical Sciences in New Delhi, says that instead of framing new rules the committee should strengthen existing guidelines "that are working reasonably well". At present, each research institute has an animal ethical committee that follows the guidelines specified in the 1960 Prevention of Cruelty to Animals Act.

Nitya Nand, former director of the Central Drug Research Institute in Lucknow, says the new rules are "most unimaginative, coming at a time when our laboratories are struggling hard to be internationally competitive". He says the committee should have just laid down guidelines for animal housing and husbandry to minimize pain and suffering: "All other functions should be left to institute ethical committees."

K.S. Jayaraman

Controversial appointment to SA council

[CAPE TOWN] Malegapuru William Makgoba, 45, a research professor at the University of the Witwatersrand, has been appointed president of South Africa's Medical Research Council (MRC) for a five-year term from next year. He replaces retiring incumbent Wally Prozesky.

Makgoba was at the centre of a fierce controversy in 1995 after allegations of incompetence, falsifying his curriculum vitae and disloyalty to his institute were made against him by 13 of the Witwatersrand's senior academics (see *Nature* 378, 324; 1995).

The former immunologist at the Royal Hammersmith Postgraduate Medical School in London became deputy vice-chancellor of the Witwatersrand in 1994, and was tipped to become vice-chancellor. But following the controversy he was transferred to a research chair in molecular immunology in 1996 for the remainder of his five-year contract with the university.

Makgoba's appointment at the MRC was



Makgoba: survived controversy at the Witwatersrand to lead the MRC.

unusual since he was chair of the MRC board. He resigned from this post after applying to be MRC president. The selection process was conducted by the deputy chair of the MRC board, Marian Jacobs, who succeeds Makgoba as chair.

Jacobs, who is head of the Child Health Unit at the University

of Cape Town, said that Makgoba had been chosen unanimously by the board after in-depth interviews and public lectures by three candidates.

Makgoba says his priorities as the MRC's president will be improving social cohesion, building capacity and promoting interdisciplinary research.

Michael Cherry

British government responds to worries over BSE in sheep

[PARIS] The British government and farming organizations moved swiftly this week to play down concerns over the risk that bovine spongiform encephalopathy (BSE) may have passed to sheep and become endemic. Their action followed publication of a news article in last week's *Nature* (395, 6–7; 1998), which reported that only nine sheep had been tested for BSE, and that the UK Department of Health had refused a proposal from the Consumers' Association to advise parents against feeding lamb to children.

In a statement, Sir Kenneth Calman, Britain's chief medical officer, said the government's Spongiform Encephalopathy Advisory Committee had advised that there was no new evidence of BSE in sheep, and concluded that no further health precautions were needed.

The health department said that "the government recognized that parents will be concerned about risks to their children, and will often choose to reduce them below a level they would accept for themselves". The Consumers' Association said that "the handling by the government of the BSE in sheep issue is exactly the situation that we wished to avoid," in that the public statements made were "unhelpful to consumers".

Ben Gill, the president of the National Farmers' Union, attacked the open discussion of the issue as "scaremongering". There is "no news here," he said, describing it as "an old story resurrected by *Nature*".

Science policy in flux in New Zealand

[SYDNEY] Science has lost political status in New Zealand following the downgrading in cabinet rank of Maurice Williamson, minister for research, science and technology, during a reshuffle. This comes at a time when the government is facing criticism from New Zealand scientists over shrinking funds and a radical plan to change priorities for ministry research grants (*Nature* 393, 198; 1998).

Williamson has also been appointed assistant to education minister Wyatt Creech, where he is expected to take over responsibility for universities. The move has led to speculation that the research ministry will be merged with the more powerful ministries of education or commerce.

Creech, who has been appointed Deputy Prime Minister, maintains control over the delayed White Paper on Tertiary Education (*Nature* 392, 320; 1998), which is expected to make controversial changes to funding formulae. Under fierce debate are whether

universities will be ranked for quality and their researchers made to compete for grants, and how research by graduate students will be supported.

UK to consider offshore wind energy investment

[LONDON] British energy minister John Battle last week announced plans to investigate developing offshore wind energy, as part of the government's drive to achieve 10 per cent of the United Kingdom's electricity needs from renewable sources by 2010.

Offshore wind is one of the United Kingdom's greatest natural and pollution-free resources, and is as yet untapped, said Battle. But he added that the contribution it could make would depend "primarily on the environmental acceptability of offshore wind projects and the technology's cost in comparison to other renewables".

Greenpeace welcomed Battle's announcement. But its own report, published during the summer, concludes that offshore wind alone could generate more than 10 per cent of Britain's electricity needs by 2010 and 40 per cent by 2030.

No need for thyroid screening, says report

[WASHINGTON] A US report released last week has recommended against widespread screening for thyroid cancers possibly caused by radioactive fallout from Cold War nuclear tests.

The report by the Institute of Medicine (IOM) and the National Academy of Sciences (NAS), *Exposure of the American People to Iodine-131 From Nevada Nuclear-Bomb Tests*, analyses a study by the National Cancer Institute (NCI) released last year, which suggested that 10,000 to 75,000 additional thyroid cancers may have resulted from above-ground tests in the 1950s (see *Nature* 389, 534; 1997).

The IOM/NAS report says widespread screening to try to identify such cancers is unwarranted because of the difficulty of measuring individual disease risk and the small number of cancers, which it put closer to the low end of the NCI's estimate. Instead it recommended public education to help those at high risk identify themselves.

Russian scientists prepare to fight for pay

[MOSCOW] Russian scientists have at last started to get their July salaries, now that about half of the promised Rubl 535 million for salaries has been transferred to scientific organizations. But August salaries are rumoured to be reduced by 15 per cent – and scientists are taking action.

Because of the fall in the ruble, the salaries

of Russian scientists are worth only a quarter of their value in September 1995, the last time salaries were raised.

The Russian committee of the scientific collectives (RCSC), which unites the trade unions of science workers, has sent a telegram to president Boris Yeltsin and the acting prime minister Viktor Chernomyrdin, drawing their attention to the "catastrophic situation" in science and demanding that salaries be paid in full.

If the RCSC gets no satisfactory reply, it plans a series of protests during September. These will culminate in a blockade of three highways leading to the capital and an all-Russia protest organized by the federation of independent trade unions on 7 October.

Five scientists killed in SwissAir 111 crash

[WASHINGTON] Jonathan Mann, founding director of the World Health Organization's Global Program on AIDS, and his wife Mary-Lou Clements-Mann, an AIDS-vaccine researcher, were among at least five scientists killed when a SwissAir flight from New York to Geneva crashed off Nova Scotia last week.

Eugenia Spanopoulou, an immunologist and Howard Hughes Medical Institute investigator at the Mount Sinai Medical Center in New York, was also on board the flight, along with her husband and child.

Two European physicists working at the Brookhaven National Laboratory were also killed in the crash — Klaus Kinder-Geiger, a nuclear physicist from Germany, and Per Spanne, a Swedish guest at Brookhaven from the European Synchrotron Radiation Facility at Grenoble in France.

NIH budget boost wins Senate committee OK

[WASHINGTON] A key committee of the US Senate last week unanimously approved a \$2 billion boost in spending on the National Institutes of Health (NIH) in 1999.

In the approved spending bill, for the Departments of Labor, Health and Human Services and Education, the NIH will receive \$15.6 billion, a 14.7 per cent increase over the 1998 level. The bill preserves the bulk of the social programmes that had been cut in a corresponding House of Representatives bill (see *Nature* 394, 108; 1998).

The bill could reach the Senate floor as soon as this week, after which it must be reconciled with a House bill that includes a 9.1 per cent NIH increase, to \$14.9 billion. Both houses must pass a final, reconciled bill before the 1999 fiscal year begins on 1 October.

The Senate Appropriations Committee's bill directs that \$175 million be spent by NIH on prostate cancer research — nearly doubling the \$89.5 million now spent.

Toxic spill caught Spain off guard

Sir — The toxic spill by the mining company Boliden into the Guadamar River and the Doñana Natural Park last April highlighted the lack of co-ordination between Spain's central and regional administrations, and their incapacity to deal efficiently with major environmental hazards. This disastrous event has once more stressed the necessity for more well-trained scientists, engineers and managers in Spain who could deal with such hazards.

The lack of trust and cooperation between the seven institutions involved in the prevention and treatment of the spill has made it painfully obvious that Spain must modernise its institutional framework for the management of water, mines and natural resources.

The Association for the Advancement of Science and Technology in Spain (AACTE), and most Spanish citizens, were shocked by the inability of official bodies to tackle the public health and environmental issues, which were worsened by their conspicuous efforts to minimize the political consequences.

One of the few positive actions was by César Nombela, president of the Spanish National Research Council (CSIC), who created a 'commission of national experts' to provide guidelines for the treatment of the toxic spill and the monitoring of its

effects. But the commission includes at least 17 members. Many are academics without much experience of detoxification, ecotoxicology and monitoring of soils, rivers and ecosystems following mine spills. The AACTE strongly suggests that the commission should be made smaller and incorporate international engineers and scientists with relevant expertise.

We urge the commission to produce a report evaluating the impact of the spill. This would include an overview of the processes following soil and groundwater pollution; effects on public health; future use of affected soils and rivers; and damage to ecosystems. Such a report would be an invaluable step towards designing detoxification and management strategies. The report should be made public to keep Spanish citizens informed and allow evaluation by the international scientific community.

We congratulate the Andalusian Environmental Council and the CSIC for setting up a monitoring programme for water and aerosols, and for allowing access to it through the Internet. The programme should be extended to soils, groundwater, river sediments, and flora and fauna. Results should be made public. This would prevent the current wave of public distrust and the occurrence of potential frauds.

The institutions involved in the

management of the area should resolve the lack of a structured decision-making system both inside and outside the Doñana Natural and National Parks. This effort must go hand-in-hand with public information, to prevent mistakes hidden behind a wall of institutional obscurity. We ask for rigorous problem analysis, professional implementation of solutions and transparent decision-making.

Many politicians still seem to be unaware of the complexity of the problem. The effects of the spill will persist for decades and require long-term action. The government must seek advice from the world's best specialists so that we can begin to learn from such calamities.

Antonio Aparicio
(President, AACTE)

*Instituto de Astrofísica de Canarias, C. Vía Láctea,
E38200 La Laguna, Tenerife, Canary Islands, Spain
e-mail: aaj@iac.es*

Javier Escartín

*Department of Geology and Geophysics,
University of Edinburgh, UK*

Luis Santamaría

*Department of Plant and Animal Interactions,
Centre for Limnology,
Netherlands Institute of Ecology, The Netherlands*

Pablo Valverde

*Department of Earth Sciences,
Memorial University of Newfoundland, Canada*

Wacky, weird and scientifically illiterate

Sir — I am surprised at your response to my comments on science and pseudo-science in the *Independent* newspaper ("How not to respond to *The X-Files*", *Nature* 394, 815; 1998). Part of the point of my piece was to warn against the too easy condemnation of popular fantasies like *The X-Files* as anti-scientific. It is heavy-handed and patronizing to try to police such stuff too thoroughly on behalf of the ideal of scientific rigour.

But to suggest, as you do, that, because science proceeds from the known to the unknown by means of a series of hypotheses, it is therefore "more like *The X-Files* than some detractors recognize" is poppycock. Science does proceed from the known to the unknown by the testing of hypotheses; but Mulder and Scully do not; rather, they parody this process in scenes whose only apparent purpose (other than to entertain) is to obfuscate and to mystify.

Where science strives to illuminate, *The X-Files* strives to darken. Even this I don't

much mind; if people enjoy obscurity, let them have obscurity. But what I do find rather irritating about the particular brand of obscurity presented in *The X-Files* is that it regularly evokes scientific concepts to help the plot, with absolutely no glimmer of understanding of what these concepts are supposed to mean. There is a scientific illiteracy in many of the scripts; and this is irritating, because there is no reason why even the wackiest and weirdest plots cannot exploit science cleverly and elegantly, rather than ignorantly and clumsily.

To complain about the clumsy use of science in drama is not, I think, to patronize anyone; but to suppose that the interaction of Mulder and Scully is "as scientific as you please", and that its popularity suggests "that the public clearly has more of a feeling for the spirit of scientific enquiry than some give it credit for", is fanciful in the extreme. There is plenty of evidence to suggest that the public is fascinated by genuine science; but the popularity of *The X-Files* must, I fear, be put down to causes other than this.

John Durant

*Imperial College London,
London SW7 5NH, UK*

Water, water, every weekend

Sir — Randall Cervený and Robert Balling's report on the raininess of weekends shows that there is nothing new under the Sun — at least not on this topic (*Nature* 394, 561–563; 1998). The same not-so-sunny conclusion was reported by David M. Schultz of the National Oceanic and Atmospheric Administration/National Severe Storms Laboratory in Norman, Oklahoma, in the *Annals of Improbable Research* (March/April 1998).

Schultz used daily weather observation records archived by the National Climatic Data Center for Atmospheric Research. A stickler for completeness, Schultz examined 40 years' worth of data, thus making his project strictly biblical in scope.

The Schultz report will be archived on our website (<http://www.improbable.com>) when last weekend's flooding subsides enough for us to re-enter our building.

Marc Abrahams (Editor)

*Annals of Improbable Research, PO Box 380853,
Cambridge, Massachusetts 02238, USA*

River meanders in a tray

Gary Parker

Meanders are a feature of some of the world's noblest rivers, but laboratory models have failed to meander convincingly. A new experiment succeeds, promising us a refined understanding of this familiar but vexing phenomenon.

In the pursuit of science it sometimes happens that an energetic amateur succeeds where specialists have repeatedly failed. Such is the case with the experimental replication of river meandering on a small scale reported by Charles E. Smith in the October issue of *Geomorphology*¹.

As seen from the air, meandering rivers offer a feast for the eye (Fig. 1). The bends describe a smooth, orderly spatial oscillation that never quite repeats itself. Although the Fly River shown in Fig. 1 appears to be frozen in time, a closer look reveals the history of its leisurely migrations across the surface of its floodplain. Repetitive rows of lacey lines on the floodplain, known as scroll bars, mark previous positions of the channel banks. Meandering rivers do not seek some ideal sinuosity or amplitude²; instead, the bends migrate laterally and downstream, interacting with their neighbours. Eventually a bend can become so tight that it is cut off and left on the floodplain as a relict oxbow lake that gradually fills with sediment. Particularly remarkable in Fig. 1 is a semicircular indentation on the upstream, outside bank of a sharp bend. The indentation, known as a concave bank

bench, has been dragged along by the migrating river bend for centuries, inscribing its half-moon signature on the floodplain.

The Fly River is an alluvial river — that is, one in which the bed and banks have been freely constructed by the interaction between flowing water and sediment. Well-established theory suggests that meandering is an instability phenomenon, so that any deviation from the straight state is self-perpetuating². And yet a crucial part of the theory, the mechanism that maintains channel coherence, is only imperfectly understood.

Meandering alluvial rivers typically have well-vegetated floodplains that are rich in fine-grained cohesive sediment. The self-reinforcing combination of fine sediments and vegetation seems to make the floodplain resistant to erosion, and helps slow down the rate of lateral erosion at the outside of a bend to the point that deposition at the opposite inside bend can keep pace. In this way the river channel can maintain its coherence as it migrates. In the absence of floodplain stabilization, a river often devolves into the less coherent (but equally interesting) braided state, as

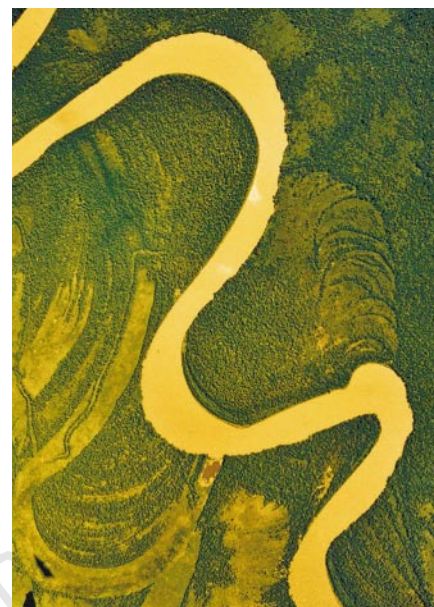


Figure 1 The meandering Fly River, Papua New Guinea (flow from top to bottom). The near absence of any human development on the floodplain means that the sedimentary structures created by channel migration can be clearly seen. Note the nearly parallel successions of long, gently curved lines known as scroll bars that mark previous channel banks, the abandoned oxbows created where bends have been cut off, and the short crescent features associated with the concave bank bench visible as the indentation in the sharp bend in the lower right of the photograph. As discussed here, Smith¹ has created tiny relatives of these characteristics of meanders in his garage. (Photograph courtesy W. E. Dietrich; channel width is on the order of 250 m.)



Figure 2 The braided Aichilik River on the north slope of the Brooks Range, Alaska, USA (flow from top to bottom). This channel form contrasts with the meanders shown in Fig. 1. (Photograph courtesy C. Paola; the width of the braided valley flat is on the order of 500 m.)

depicted in Fig. 2. Indeed, Murray and Paola³ have proposed that braiding is the default morphology of alluvial streams, and that meandering occurs only when braiding tendencies are suppressed.

Many researchers have attempted to model meandering in the laboratory at reduced scale. The bends form readily and spontaneously, only to be followed by channel widening and a devolution to the braided state before the bends achieve appreciable sinuosity. Until now, the best attempt⁴ involved the use of clay as an agent to stabilize sediment deposits on the inside of bends, but even in this case the bends were only mildly sinuous.

Smith is a software specialist who devotes his spare time to the modelling of meanders in a 'river tray' in his garage. His meandering channels have a width of only about 4 cm. They evolve from a straight alignment to channels that are as sinuous as any seen in nature. Concave bank benches, bend cutoffs and features analogous to scroll bars form readily in spite of the vast difference in scale and the fact that the flow is barely turbulent; this suggests that the

role of turbulence in meander mechanics is less critical than previously supposed. Although bend shape is somewhat different from that seen in the field, the experimental meanders are clearly close relatives of field meanders. The key to the success of the experiments seems to be the use of a mixture of diatomaceous earth (a natural sediment rich in diatom skeletons) and kaolinite clay as the model sediment. This mixture provides just the degree of cohesion that allows stabilization of the deposits on the inside of bends while slowing but not preventing bank erosion.

Further study of these small bends should tell us more about how a much larger meandering stream, such as the Fly River, maintains its integrity as a highly sinuous, single-thread channel as it wanders across its floodplains. In existing theoretical treatments of meandering, deposition at the inner bank is simply assumed to keep pace with erosion at the outer bank of a bend, so as to maintain

constant width. Smith's technique offers for the first time an experimental analogue that essentially reproduces this behaviour of its own accord. It thus opens a window into study of the interaction of riverbank dynamics with meander dynamics.

Now, in the mind's eye sprinkle the floodplain of the Fly River with the various engineering works of human civilization — bridges, ports, roads, pipeline crossings — that are constructed under the misunderstanding that the river does not move. If Smith's approach can teach us how rivers move, it can help the engineer better design for and live with that eventuality.

Gary Parker is at the St Anthony Falls Laboratory, University of Minnesota, Mississippi River at Third Avenue, Minneapolis, Minnesota 55414, USA. e-mail: parke002@tc.umn.edu

1. Smith, C. E. *Geomorphology* **25**, 19–30 (1998).
2. Parker, G. & Andrews, E. D. *J. Fluid Mech.* **162**, 139–156 (1986).
3. Murray, A. B. & Paola, C. *Nature* **371**, 54–57 (1994).
4. Jin, D. & Schumm, S. A. in *1st Int. Geomorphology Conf.* (ed. Richards, K. S.) 680–691 (Wiley, Chichester, 1986).

Developmental biology

Casting an eye over cyclopia

Patrick Blader and Uwe Strähle

In ancient Greek mythology the cyclops had a single, central eye. Today, cyclopia is an extreme characteristic of a class of human birth-defect syndromes known as holoprosencephaly. Although the morphological abnormalities that underlie holoprosencephaly are well characterized — a lack of, or defect in, development of the ventral brain — the biochemical defects that underlie this malformation have remained elusive. But the analysis of cyclopic zebrafish (*Danio rerio*) mutants by Feldman *et al.*¹ and Sampath *et al.*² (on pages 181 and 185 of this issue), and by Rebagliati *et al.*³ in *Proceedings of the*

National Academy of Sciences, provide an insight into the mechanisms that establish the ventral regions of the brain and spinal cord.

Cells in the mesoderm layer at the midline of the embryo secrete signals that establish midline identity in the overlying neural plate. Among these signals are members of the hedgehog (Hh) family of proteins, which can induce a variety of cells, including the neural midline. This region gives rise to the ventral brain of the embryo, and the floorplate and motor neurons in the spinal cord⁴. Mouse embryos with mutations in the gene that

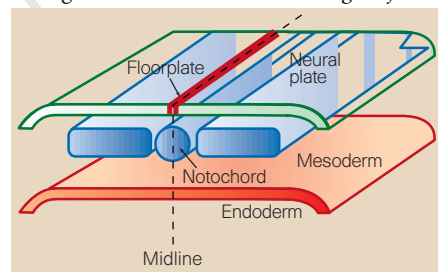


Figure 1 Structures formed during embryogenesis. Early embryogenesis is characterized by a series of morphogenetic movements, known as gastrulation, which establishes the three primary germ layers — the ectoderm, mesoderm and endoderm. At the end of this process, the mesoderm (the future muscles and bones) comes to rest sandwiched between the ectoderm (the future nervous system) on the outside of the embryo, and the endoderm (the future gut) on the inside.

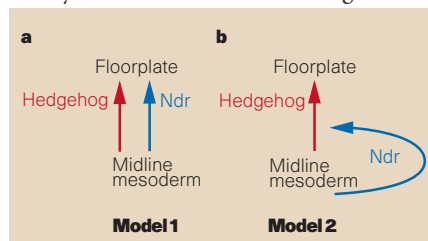


Figure 2 Alternative models of how nodal and nodal-related factors (Ndr) may act in specification of the floorplate. a, Ndr secreted from notochord and prechordal plate cells act directly in combination with proteins of the hedgehog family on floorplate precursor cells. b, Ndr act indirectly in floorplate specification. The function of the Ndr is confined to the axial mesoderm (notochord and prechordal plate), where they may be involved in processing of hedgehog signals or in controlling the expression of additional, as-yet-unidentified signals.

encodes Sonic hedgehog (Shh) do not develop midline structures of the neural plate, leading, at worst, to profound cyclopia. A single, medial eye is located above a nose containing a single nostril⁵. But although Shh is necessary for specification of the neural plate, other results indicate that it may not be enough. Cloning of the genes responsible for two cyclopic zebrafish mutant phenotypes now suggests that members of the transforming growth factor- β (TGF- β) family may also be involved in these processes.

The TGF- β superfamily is a large group of secreted polypeptide growth factors. Mouse embryos containing two mutated copies of the TGF- β -related gene *nodal* cannot form mesoderm⁶. Likewise, mice with mutations in the *Smad2* gene, which is predicted to transduce signals downstream of an as-yet-unidentified nodal receptor, gastrulate abnormally and lack mesoderm (Fig. 1)^{7,8}. In a particularly insightful experiment, Nomura and Li⁷ previously analysed mice that were heterozygous for mutations in both *nodal* and *Smad2*. One-third of the embryos had profound cyclopia, suggesting that nodal signalling may also be required to specify the midline of the neural plate. But for more clues about this, we must shift to the mutants identified in the zebrafish.

Zebrafish embryos with mutations in the *cyclops* (*cyc*) gene have defects in specification of the ventral brain and floorplate. As the name suggests, these defects result in a cyclopic phenotype⁹. Sampath *et al.*² and Rebagliati *et al.*³ now show that the *cyc* gene encodes a TGF- β signalling molecule with 69% identity to the mouse nodal protein⁶. Moreover, Feldman *et al.*¹ show that a second, milder cyclopic mutation, *squint* (*sqt*), is an enhancer of the *cyc* phenotype. *sqt/cyc* double mutants lack mesoderm. *sqt* encodes another nodal-related protein¹. As with mouse *nodal*, *cyc* and *sqt* are expressed at the correct place and time, in keeping with a possible role in controlling the fates of midline cells in the neural plate^{1,2,10}. Interestingly, *cyc* mutants lack the floorplate but develop motor neurons⁹, suggesting that the Shh signalling pathway is still active. Because mutations in either the Shh or nodal signalling pathways lead to defects in the ventral neural tube, perhaps nodal modulates the target-cell response to Shh signals. Thus, whereas both *Cyc* and *Shh* would be required to induce the floorplate, *Shh* alone would be enough to induce motor neurons.

However appealing this model is, it leaves some questions. The evidence that a nodal-like signal needs to be received by the neural plate for induction of the floorplate and ventral-brain fates remains circumstantial. Mesoderm at the anterior

end of the midline, known as the prechordal plate mesoderm (PCP), is affected in both *cyc*¹¹ and *sqt*¹² single mutants, suggesting that the functions of these genes do not overlap in maintaining the PCP. It is possible, therefore, that *cyc* and *sqt* mutations lead indirectly to cyclopia through a loss of Shh signalling by the PCP (Fig. 2). Moreover, although Shh is expressed in the notochord (the midline mesoderm posterior to the prechordal plate) of *cyc* mutants, Cyc may, nonetheless, be required for processing and secretion of Shh by adjacent notochordal cells.

Sampath *et al.*² have also reopened the question of when the floorplate is specified during development. When they forced expression of Cyc in the mature notochord of *cyc* mutants, the floorplate in the immediately overlying neuroectoderm was not rescued. This suggests that floorplate specification occurs during the early stages of gastrulation, when *cyc* is first expressed, and when future floorplate and notochordal cells are intermingled.

Despite these unresolved issues, the three zebrafish reports^{1–3} provide an insight into the function of nodal-related signals during formation of the germ layers and development of the axis. There is also a more general implication — two *nodal*-related genes in the zebrafish must be mutated to produce the phenotype of the single *nodal* mutation in the mouse. Together with evidence¹³ that the genomes of teleosts

(bony fishes) have duplicated during evolution, this might give pessimists among us ammunition against using the zebrafish as a genetic system to study development. But it is precisely the partial redundancy of *cyc* and *sqt* that has allowed the dissection of both an early function for these genes, in induction of the mesoderm, and a late function in neural-plate patterning — events that are still unclear in the mouse system. It is to be hoped that future analysis of combinations of zebrafish mutants will provide us with similar insights into the mechanisms of embryogenesis. □

Patrick Blader and Uwe Strähle are in the Institut de Génétique et de la Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, BP163, 67404 Illkirch Cedex, CU de Strasbourg, France.
e-mails: blader@igbmc.u-strasbg.fr
uwe@titus.u-strasbg.fr

1. Feldman, B. *et al.* *Nature* **395**, 181–185 (1998).
2. Sampath, K. *et al.* *Nature* **395**, 185–189 (1998).
3. Rebagliati, M. R., Toyama, R., Haffter, P. & Dawid, I. B. *Proc. Natl Acad. Sci. USA* **95**, 9932–9937 (1998).
4. Tanabe, Y. & Jessell, T. M. *Science* **274**, 1115–1123 (1996).
5. Chiang, C. *et al.* *Nature* **383**, 407–413 (1996).
6. Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. & Kuehn, M. R. *Nature* **361**, 543–547 (1993).
7. Nomura, M. & Li, E. *Nature* **393**, 786–790 (1998).
8. Waldrip, W. R., Bikoff, E. K., Hoodless, P. A., Wrana, J. L. & Robertson, E. J. *Cell* **92**, 797–808 (1998).
9. Hatta, K., Kimmel, B. C., Ho, R. K. & Walker, C. *Nature* **350**, 339–341 (1991).
10. Rebagliati, M. R., Toyama, R., Fricke, C., Haffter, P. & Dawid, I. B. *Dev. Biol.* **199**, 261–272 (1998).
11. Thisse, C., Thisse, B., Halpern, M. E. & Postlethwait, J. H. *Dev. Biol.* **164**, 420–429 (1994).
12. Heisenberg, C. P. & Nüsslein-Volhard, C. *Dev. Biol.* **184**, 85–94 (1997).
13. Postlethwait, J. H. *et al.* *Nature Genet.* **18**, 345–349 (1998).

sediments from the *Challenger* expedition, and they have also been found in polar ice and sometimes in sedimentary rocks. The sediment samples are particularly intriguing because they provide a historical record of the infall of meteoric bodies, and they also provide information on atmospheric composition. Spherules with terrestrial ages as great as 190 million years have been reported⁴, but they are heavily degraded by chemical weathering and until now there was little hope that the historical record could be extended further back in time.

The atmospheric record is imprinted on cosmic spherules as they fall. Spherules form at altitudes above 80 km, where frictional heating of small meteoroids causes them to flash melt at temperatures above 1,500 °C. Both the stony and iron types of spherule interact strongly with the atmosphere, incorporating terrestrial oxygen during the few seconds that they are molten. During this time, they encounter a mass of air a few times their own mass. With the present atmospheric oxygen content, that is just enough to totally oxidize the iron–nickel spherules, so those we find are largely composed of the oxides magnetite and wüstite, with only small residual metallic cores of iron and nickel or of platinum-group elements. The stony spheres contain both pre-terrestrial and atmospheric oxygen, and their final oxidation state is determined by the oxygen content of the atmosphere.

In earlier times, we believe that the atmospheric oxygen abundance was lower. The rise of oxygen in the atmosphere was a momentous event in Earth's history because it allowed the advance of aerobic metabolism and the explosive development of life during the Cambrian period, some 500 to 540 million years ago. But the initial rise of oxygen to moderate levels is believed to have occurred between one and two billion years ago, and the discovery of preserved cosmic spherules from this period provides a new means to detect the rise. This intriguing means of estimating past oxygen partial pressure would be a true global measurement, because the oxidation of the spherules occurs in a

Planetary science

Ancient cosmic spherules

Don Brownlee

On page 146 of this issue¹, Deutsch, Greshake, Pesonen and Pihlaja report the spectacular discovery of cosmic spherules in sandstone from Finland that is 1.4 billion years old (Fig. 1). These spherules hit the Earth far earlier than any other meteoritic material we have, and yet they are in an excellent state of preservation. Their existence provides a new means of tracing long-term changes in the oxygen content of the atmosphere, as well as changes in the Solar System's meteoroid population.

In mid-November, the Earth will pass through the orbit of Comet Temple–Tuttle, and the annual Leonid meteor shower may once again repeat the sensational 'meteor storm' displays observed² in 1799, 1833, 1866 and 1966. But whether or not the 1998 Leonids are a great storm or just a typical shower, they will be a vivid reminder that the Earth is incessantly bombarded by objects from space. Nearly 40,000 tonnes of comet and asteroid debris hit the Earth each year, most of it as tiny particles about 200 micrometres in diameter. All of them enter

the atmosphere at high speed and some vaporize, but others survive, either as micrometeorites or as melted particles called 'cosmic spherules'.

Cosmic spherules are the most abundant form of extraterrestrial material on Earth. They were first described in 1876 by Sir John Murray³, who found them in deep ocean



Figure 1 Small, but perfectly formed. This cosmic spherule, a fraction of a millimetre across, was formed when a dust particle hit the Earth's upper atmosphere 1.4 billion years ago and melted. Remarkably, it and its contemporaries have been preserved unweathered in Finnish sandstone.

well-mixed layer of the atmosphere.

These old cosmic spherules also provide a way of detecting long-term changes in the meteoroid environment of the Solar System. Infall rates over the past 1,000 years have been measured by counting spherules in ice melted at the South Pole⁵. For older samples, where it may be difficult to recover individual spherules, measuring helium and osmium isotopes in the bulk rock are a handle on long-term change^{6,7}.

There is a strong link between these laboratory studies of Solar System dust and astronomical studies of dust around other stars. That is of strong current interest: in a poster at the July 1998 Protostars and Planets IV meeting, C. Dominik *et al.* presented data from the Infrared Space Observatory indicating that up to 50% of nearby young stars are surrounded by detectable dust envelopes. It seems likely that many of these envelopes are maintained by comets or asteroids, implying the presence of planets or at least the building blocks of planets.

Understanding circumstellar dust systems is also crucial for the design and implementation of space-based interferometric telescopes, such as the Darwin and Terrestrial Planet Finder missions, currently being designed to detect extra-solar terrestrial planets. These missions must be able to distinguish planets from the large infrared background caused by circumstellar dust, and so the properties, distribution and evolutionary history of dust in our own Solar System will provide important 'ground truth' for other stars. □

Don Brownlee is in the Department of Astronomy, Box 351580, University of Washington, Seattle, Washington 98195, USA.

e-mail: brownlee@astro.washington.edu

1. Deutsch, A., Greshake, A., Pesonen, L. J. & Pihlaja, P. *Nature* **395**, 146–148 (1998).
2. <http://www-space.arc.nasa.gov/~leonid/index.html>
3. Murray, J. C. *Proc. R. Soc. Edinb.* **9**, 258 (1876).
4. Czajkowski, J. *Meteoritics* **18**, 286–287 (1983).
5. Taylor, S., Lever, J. H. & Harvey, R. P. *Nature* **392**, 899–902 (1991).
6. Farley, K. A. *Nature* **376**, 153–156 (1995).
7. Schmitz, B., Peucker-Ehrenbrink, B., Lindstrom, M. & Tassinari, M. *Science* **278**, 88–90 (1997).

Population ecology

Common rules for animals and plants

John D. Damuth

The causes of the apparent regularities in the relationship between plant size and number in natural stands of vegetation¹ constitute something of a puzzle. It is no surprise that, as plants in a population grow larger, a given area can support fewer of them. Likewise, a flora dominated by large plants has a lower density of such plants than does vegetation dominated by smaller species. But what determines these dynamics? Plant ecologists have long thought that the controlling factors in size–density scaling would be the geometric properties of individual plants. Geometry-based theory suggests that average plant size should scale as the $-3/2$ power of population density. This is known as the 'thinning law', and there has been a lively controversy about whether empirical observations support it.

On page 163 of this issue², Enquist *et al.* take a different approach to the thinning law, based on the rate of energy use of individual plants. They predict a $-4/3$ exponent for the thinning relationship and show that many data are in good agreement with the prediction. The same exponent applies to comparable relationships in animals. So, remarkably, it appears that among plants both metabolism and abundance scale allometrically in much the same way as they do in animals, although in neither case do we yet understand why.

Actually, there are two distinct size–density relationships under consideration³.

The intraspecific relationship (the only one properly termed a 'thinning' relationship) is a trajectory observed over time in populations of a single species, among growing individuals subject to density-dependent mortality. In contrast, the interspecific relationship concerns the scaling of plant size and number across stands of species that differ in average size at maturity. These two relationships are often discussed together because they usually have similar scaling exponents, but in fact they need not be closely related either mathematically or causally.

The traditional geometric model for plant thinning is straightforward. The area of ground that a plant covers is assumed to scale as the square of the diameter of the plant's canopy footprint, and the plant's volume to scale as the cube of this diameter. So volume scales as the $3/2$ power of coverage area. The number of individuals that can fit in a given area will be the reciprocal of coverage area. Volume (and hence plant mass) should therefore be proportional to the $-3/2$ power of population density. This will describe the scaling of maximum plant size once the population has become dense enough for crowding to affect survival.

Unfortunately, there is no unequivocal evidence for such a relationship³. Observed thinning exponents are highly variable among species, and attempts to adjust the

expected exponent by elaborating more specific geometric models have had only limited success.

But what if something other than physical crowding is the main influence on the number of individuals a given area can support? In particular, ask Enquist *et al.*², what if plants merely grow until they are limited by their resources? To derive a thinning law from this perspective, we need to know how much energy plants of different sizes require. A notable contribution of Enquist *et al.* is their empirical demonstration that total plant resource use (indirectly estimated as whole-plant xylem transport) scales both within and among species as the $3/4$ power of body mass, which is the same exponent as for animal metabolic requirements. Using this as an intraspecific value, Enquist *et al.* predict a thinning exponent of $-4/3$ if resource levels are constant.

Curiously, a thinning law for animal populations has been proposed⁴ based on exactly the same model. Because metabolism scales the same way as in plants, here too the expected exponent is $-4/3$. Lattó⁵ argues that variation in observed thinning trajectories for animal populations should be expected because, within species, metabolic rate may often scale differently from the way it does among species. This suggests a similar, non-geometrical explanation for some of the variation among plant thinning trajectories. Enquist *et al.* do not claim that their model actually predicts thinning trajectories for particular species. But this thinning model common to plants and animals begs to be investigated further.

Turning to the interspecific relationship, the results of Enquist *et al.* point to an even more intriguing similarity between abundance scaling in plants and animals. Here, the interest lies in the relative abilities of species of different sizes to reach and maintain population density as stand dominants. If the issue is physical crowding — how many individuals of a species of given size can be packed into an area — we would expect that the interspecific relationship would resemble the theoretical, geometry-based thinning relationship (exponent of $-3/2$). If, on the other hand, rates of resource use determine the relationship, we would expect something different. A variety of empirical studies, including the new analyses reported here², show that across 12 orders of magnitude in plant size the exponent actually tends to be closer to $-4/3$.

This is significant because it is equivalent (by simply rearranging terms) to saying that, across habitats, plant density scales as the $-3/4$ power of body mass. This is identical to the way population density scales in animals, over comparable ranges



100 YEARS AGO

Readers of prospectuses of educational institutions and polytechnics may have noticed that of late years there has been a tendency to convert the teachers into professors. The nature of the institution in which the instructors can rightly use the latter title is apparently a matter of opinion, and it is becoming worth while to define the duties and position of a professor. Miss Catherine Dodd describes in the *National Review* how she asked 105 primary school children, between the ages of ten and fourteen, to give this definition, among others. Here are some of the attempts: – “A man who has passed a high examination.” “A very clever man.” “One who can do his work easily.” “A man skilled in sense.” That a professor has a certain social standing is evident from the definitions which describe him as “a man who is well off,” and “a man who lives in a nice house.” Among the vague definitions are the following: – “A person who professes to do something.” “A man who says he can do anything.” But the children’s general idea is that a professor teaches music, dancing, or languages, or performs conjuring tricks. Thus, “A professor teaches all kinds of instruments.” “He is a gentlemen that generally plays at balls,” and “a man who knows clever tricks.” To correctly define a professor would probably prove a difficulty to many children of older growth.

From *Nature* 8 September 1898.

50 YEARS AGO

In the editorial article in *Nature* of August 14, in criticizing a recent statement of the Atomic Scientists’ Association, it is argued that collaboration between scientific men east and west of the ‘Iron Curtain’ may be undesirable, because it is likely “to promote, for the present, a one-way traffic to the disadvantage of the Western democracies” ... It is stated in the editorial that the man of science in totalitarian countries is essentially a servant of the State, and that it is treason for him to divulge any knowledge save as the State allows. But in fact this statement is true only in the opinions of the men who control the Government of the U.S.S.R.; we may be sure that most scientific men in the satellite countries would not take that view of their functions. – N. F. Mott.

From *Nature* 11 September 1948.

in body size⁶ — a relationship that is known as ‘energy equivalence’ because the $-3/4$ exponent indicates that population energy-use per unit area is independent of body mass.

Enquist *et al.* have shown that plant species of all sizes can achieve the same rates of local resource use. The energy equivalence relationship may thus be one of the most widespread of ecological regularities, yet we have no satisfactory explanation of its origin and maintenance. What prevents evolution from producing species that routinely violate it? Now that the relationship has been identified in plants as well as animals, the answer

becomes that much more interesting. It also suggests that the explanation (if there is a single one) must be very general. □

John D. Damuth is in the Department of Ecology, Evolution and Marine Biology, University of California at Santa Barbara, Santa Barbara, California 93106, USA.

e-mail: damuth@lifesci.ucsb.edu

1. Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. *J. Biol. Osaka City Univ.* **14**, 107–129 (1963).
2. Enquist, B. J., Brown, J. H. & West, G. B. *Nature* **395**, 163–165 (1998).
3. Lonsdale, W. M. *Ecology* **71**, 1373–1388 (1990).
4. Begon, M., Firbank, L. & Wall, R. *Oikos* **46**, 122–124 (1986).
5. Latto, J. *Oikos* **69**, 531–534 (1994).
6. Damuth, J. *Biol. J. Linn. Soc.* **31**, 193–246 (1987).

Mathematics

Nonlinear modes of vibration

Ivar Ekeland

One of the most fundamental results in classical mechanics is that linear systems with n degrees of freedom have n fundamental modes of vibration, and that any motion of the systems can be obtained as a linear combination of these fundamental modes. This combination principle does not hold for nonlinear systems, but one suspects that a few periodic solutions will play a part akin to the fundamental modes. In one case, that suspicion has now been confirmed: H. Hofer, K. Wysocki and E. Zehnder¹ have proved that so-called convex systems with two degrees of freedom must have either two periodic orbits for a given energy, or infinitely many — they cannot have three, or twenty thousand, or a billion.

To appreciate the meaning of this result, begin by considering linear systems with two degrees of freedom, such as the weight fixed to springs in Fig. 1. They can always be described in terms of two harmonic oscillators, with periods T_1 and T_2 . If the two periods have a common multiple, that is, if $aT_1 = bT_2 = T$ for some pair of integers a and b , then every motion is periodic, with period T . But if there is no common period then there are exactly two periodic motions, with periods T_1 and T_2 , in which one of the oscillators is active and the other is shut down. All other motions arise from superimposing both oscillators, and give rise to the well-known Lissajous curves, which are aperiodic.

You should try to visualize this in its four-dimensional phase space (p_1, p_2, q_1, q_2) , where p is momentum and q is position. The potential energy of the system is $V = 2\pi^2(q_1^2/T_1^2 + q_2^2/T_2^2)$, and its total energy, or Hamiltonian, is $H = 1/2(p_1^2 + p_2^2) + V$. The surfaces of constant energy are three-dimensional ellipsoids, around which wind the trajectories of the motion. If T_1/T_2 is rational, all these trajectories are closed, corresponding

to periodic motions. If T_1/T_2 is irrational, then there are only two closed trajectories for any energy, corresponding to the two fundamental modes of vibration; all other trajectories are open, filling the whole constant-energy surface. This picture extends simply to any number of degrees of freedom.

But how does it extend to nonlinear systems²? Consider a convex potential V defined in n -dimensional Euclidean (flat) space, such that V goes to infinity at infinity. In phase space, the Hamiltonian is still $H = 1/2(p_1^2 + p_2^2) + V$, and the constant-energy surfaces are convex and bounded. The corresponding mechanical system is a

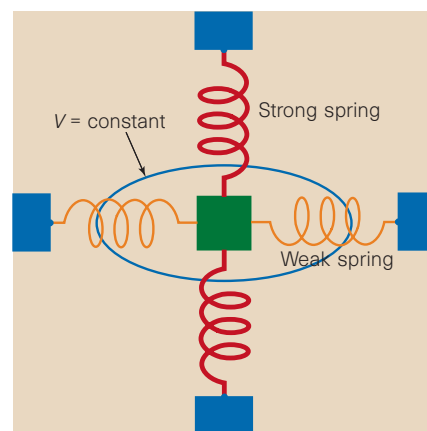


Figure 1 A dynamical system with two degrees of freedom. A mass is held by springs; strong ones in one direction, weak in the other. It can oscillate up and down rapidly, or side to side slowly, or in more complicated two-dimensional patterns. If the force in the springs depends linearly on their displacement, then lines of constant potential energy V are ellipses, and all the trajectories are well understood. The nonlinear problem is more difficult, but it has now been proved for a large class of systems that there must either be two periodic trajectories, or an infinity.

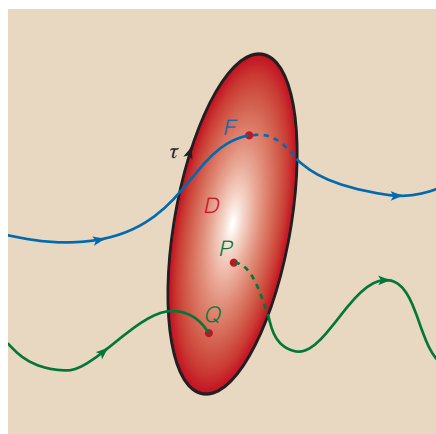


Figure 2 Trajectories in phase space. For a convex nonlinear system with two degrees of freedom, all trajectories pass through a disk, D . The trajectory leaving point F returns to the same point (perhaps after crossing the disk several times), so it is periodic; but the one leaving point P returns to Q , and will continue to meander through the allowed region of phase space aperiodically.

system of n linked nonlinear springs.

It has already been proved that, for nonlinear systems of this kind, every energy has at least two periodic solutions. Drawing a parallel with the linear case, we might guess that a nonlinear system with n degrees of freedom would carry at least n periodic solutions. That is unproved, but might well be true; it is what mathematicians call a conjecture.

This conjecture raises a question about the global behaviour of nonlinear systems. Much of what is known about nonlinear systems arises from perturbation theory, and gives results valid only for small values of some parameter. There is nothing of the kind here — instead we are asking whether we can build some specific nonlinear system that has only so many periodic solutions; in other words, we have to understand the dynamics not locally, but on the whole of the energy level. Unfortunately, systems whose trajectories can be written explicitly in terms of the initial data and elapsed time are rare. So to prove results about general nonlinear systems one cannot make explicit computations but must resort to other, more geometrical, methods.

As I mentioned before, fixing the energy level to some constant yields a three-dimensional hypersurface that is followed by any trajectory. Hofer *et al.*¹ then find, by variational methods, the periodic trajectory τ that is in some sense the simplest (in the linear case, it would be the one with the shortest period). Then they show using the theory of partial differential equations that τ defines the boundary of a two-dimensional disk D through which *all* trajectories must pass (Fig. 2). If, for instance, P is a point on the disk D itself, the trajectory issuing from P

will hit D at some other point Q , thereby defining a map from D into itself. This map turns out to be area-preserving, and it is known that such a map must have either one fixed point (corresponding to one more periodic orbit of the flow) or infinitely many.

Clearly, the proof does not carry over to higher dimensions: if there are more than two degrees of freedom, then the energy level will be five-dimensional or more; and although we could still find a periodic trajectory with much the same variational properties as τ , the two-dimensional disk spanning this closed orbit will not catch all the remaining trajectories because they have too many dimensions to move in. So the corresponding conjecture, that every convex Hamiltonian system with n degrees of freedom must carry n periodic orbits or infinitely many, remains open.

To show again that all this is far from

obvious, let us ask another question: what systems have n periodic orbits only? One example is the linear case of harmonic oscillators with no common period. Is there anything else? Michael Herman (personal communication) has shown that there is: for any n , there exists a convex system with n degrees of freedom that has exactly n periodic orbits on a prescribed energy level, and which cannot be decomposed into n independent harmonic oscillators by any change of variables. In spite of their nonlinearity, general convex Hamiltonian systems still preserve some of the features of harmonic oscillators. □

Ivar Ekeland is at the University of Paris-Dauphine, 75775 Paris Cedex 16, France.

e-mail: ivar.ekeland@dauphine.fr

1. Hofer, H., Wysocki, K. & Zehnder, E. *Ann. Math.* (in the press).

2. Ekeland, I. *Convexity Methods in Hamiltonian Mechanics* (Springer, Berlin, 1990).

HIV

Setting death in motion

Jean Claude Ameisen

Cell motion, an essential component of embryonic development, continues to be important throughout life. The immune system constantly patrols our bodies, checking for infectious pathogens then regrouping cells to attack at sites of invasion. This complex choreography is controlled by a range of chemokines (and chemokine receptors), which induce and guide the movement of cells along chemokine concentration gradients¹. But, as reported by Herbein *et al.*² (on page 189 of this issue) and Hesselgesser *et al.*³ (in *Current Biology*), chemokines may also act as death signals. And subversion of such signals by the human immunodeficiency virus (HIV) may participate in the development of disease.

An intriguing feature of HIV pathogenesis is this. How does the virus induce programmed cell death (apoptosis) not only in the population of T cells that bears the CD4 receptor (which is required for viral entry), but also in other cell types? For example, HIV causes the death of CD8⁺ antiviral effector T cells in the immune system, and neurons in the brain, leading to immune exhaustion and dementia, respectively⁴.

We know that infection by HIV-1 requires binding of its surface envelope protein, Env, to both the CD4 receptor and a chemokine receptor¹. Strains of HIV-1 that are prevalent during the first years of asymptomatic infection bind to members of the CCR chemokine-receptor family, whereas strains that become prevalent at the onset of disease bind the CXCR4 chemokine receptor. Thus, chemokines may have a beneficial effect by acting as

natural antagonists of HIV infection in CD4⁺ cells. But, unlike CD4 molecules, chemokine receptors are widely expressed in both the immune system and the brain.

Herbein *et al.*² and Hesselgesser *et al.*³ now show that the Env protein from X4 strains of HIV-1 (that bind CXCR4), and the natural CXCR4 ligand stromal-derived factor-1 (SDF-1), can induce apoptosis *in vitro* in CD8⁺ T cells² and in a neuronal cell line³. Moreover, Herbein *et al.* reveal that the death pathway downstream of CXCR4 signalling in CD8⁺ T cells is a complex and dynamic process. It involves crosstalk with another population of immune cells, the macrophages, and operates through the tumour-necrosis factor- α (TNF- α)/TNF-receptor II (TNFR II) death-transducing pathway (Fig. 1, overleaf).

Signalling by CXCR4 induces the surface expression of TNF- α in macrophages, and TNFR II in CD8⁺ T cells. Subsequent contact between the macrophages and T cells then triggers T-cell death. Although Herbein *et al.* detected only a low percentage of apoptotic CD8⁺ T cells, cell loss was high — most of the apoptotic corpses were probably ingested by macrophages, rapidly clearing the scene⁵. Interestingly, this Env-induced death of CD8⁺ T cells through CXCR4 and then TNF- α /TNFR II signalling, resembles the Env-induced death of CD4⁺ T cells through CD4 and then Fas ligand (FasL)/Fas receptor signalling⁶. Death of CD4⁺ T cells also involves the (optional) recruitment of macrophages⁷. Moreover, TNF- α /TNFR II and FasL/Fas belong to the same family of death-transducing ligand/receptor pairs,

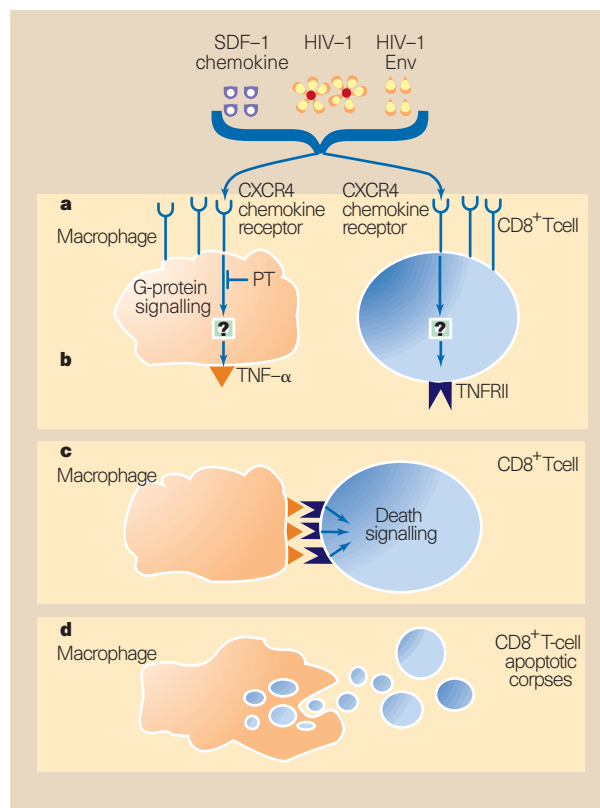


Figure 1 Chemokine receptor-mediated cell death. a, The envelope (Env) surface protein of some strains of HIV-1 can, like the stromal-derived factor-1 (SDF-1), signal through the CXCR4 chemokine receptor. b, In macrophages, CXCR4 signalling (through a pathway that can be inhibited by pertussis toxin; PT) induces surface expression of tumour-necrosis factor- α (TNF- α) and, in CD8 $^{+}$ antiviral effector T cells, of the TNF- α receptor II (TNFRII). c, Contacts between both cells trigger TNFRII-mediated death signals in CD8 $^{+}$ T cells. d, Macrophages probably clear the scene rapidly.

Finally, when they emerge, to what extent do X4 strains really substitute for SDF-1 signalling? Herbein *et al.*² show that X4 Env, just like SDF-1, triggers CXCR4 calcium signalling. Moreover, pertussis toxin, which blocks CXCR4 signalling, blocks the macrophage-mediated, TNF- α /TNFRII-dependent, CD8 $^{+}$ T-cell death. But what about cell motion? Env protein from a strain of HIV-1 that uses CC chemokine receptors can partly mimic these chemokines by attracting CD4 $^{+}$ (but not CD8 $^{+}$) T cells¹³. Does Env from X4 strains, like SDF-1, attract CD8 $^{+}$ T cells? This could explain the late appearance of X4 strains in HIV-1 patients. By attracting antiviral effectors to the site of viral production, the virus would be selected against until the loss of CD4 $^{+}$ T cells deprived the site of signals for survival of the CD8 $^{+}$ cells. But in some (rare) patients, X4 strains emerge early. HIV-1 Env is highly variable, and it may have several ways of subverting CXCR4 signalling. So it is possible that some strains of X4 mimic SDF-1, setting the CD8 $^{+}$ T cells in motion, with death (or survival) awaiting on arrival, whereas other strains of X4 signal for death instead of motion. □

Jean Claude Ameisen is at INSERM C/JF 9707, Hôpital Bichat/Université Paris 7, 75788 Paris Cedex 18, France.

e-mail: ameisen@caramail.com

1. Ward, S. G., Bacon, K. & Westwick, J. *Immunity* **9**, 1–11 (1998).
2. Herbein, G. *et al. Nature* **395**, 189–194 (1998).
3. Hesselgesser, J. *et al. Curr. Biol.* **8**, 595–598 (1998).
4. Ameisen, J. C., Estaquier, J., Idziorek, T. & De Bels, F. *Trends Cell Biol.* **5**, 27–32 (1995).
5. Savill, J. *Nature* **392**, 442–443 (1998).
6. Algeciras, A., Dockrell, D., Lynch, D. & Paya, C. *J. Exp. Med.* **187**, 711–720 (1998).
7. Badley, A. *et al. J. Exp. Med.* **185**, 55–64 (1997).
8. Zheng, L. *et al. Nature* **377**, 348–351 (1995).
9. Rosenberg, E. S. *et al. Science* **278**, 1447–1450 (1997).
10. Zou, Y. R., Kottmann, A. H., Kuroda, M., Taniuchi, I. & Littman, D. R. *Nature* **393**, 595–599 (1998).
11. Kledal, T. N. *et al. Science* **277**, 1656–1659 (1997).
12. Hadida, F. *et al. J. Exp. Med.* **188**, 609–614 (1998).
13. Weissman, D. *et al. Nature* **389**, 981–985 (1997).

participating in the physiological control of CD8 $^{+}$ and CD4 $^{+}$ T-cell survival, respectively⁸. Thus, Env from X4 strains of HIV-1 may induce premature death in both populations of T cells, by subverting the very signals that usually downregulate successful immune responses.

The finding that the chemokine SDF-1 can also act as a death signal raises several questions. SDF-1 blocks infection of CD4 $^{+}$ T cells by X4 strains¹, so, during the first years of HIV infection, it may act as a double-edged sword — preventing the emergence of X4 strains, while initiating part of the disease. But once the cellular source of SDF-1 is exhausted, high levels of X4 will accelerate the course towards disease. Such SDF-1-mediated damage could explain why half of the patients in whom X4 never emerges can still progress (at a slower pace) towards disease.

But why should chemokines behave as death signals in the first place? By coupling the regulation of programmed cell death to that of cell differentiation, proliferation and migration, a stringent 'social control' can be placed on cell behaviour. Because CD8 $^{+}$ T cells are professional killers, they may do serious harm if they end up in the wrong place. So perhaps chemokines such as SDF-1 set CD8 $^{+}$ T cells on a journey, with macrophage-mediated death on arrival, unless appropriate survival signals ensure that they have reached the appropriate location. There, competition for limited amounts of such signals may continuously select for survival of the fittest antiviral effectors. Survival signals may be provided,

directly or indirectly, by the CD4 $^{+}$ T cells, which coordinate the immune response. Indeed, efficient immune control of HIV *in vivo* has been linked to CD4 $^{+}$ T-cell functions⁹, such that when patients lose these functions, SDF-1 may summon CD8 $^{+}$ T cells towards death. Because CXCR4 and SDF-1 are vital to embryonic development¹⁰, therapeutic strategies aimed at blocking CXCR4 signalling¹¹ will depend on the importance of such signalling in adult body function. Alternatively, increasing CCR signalling may be useful, because some CC chemokines, in contrast to SDF-1, increase the CD8 $^{+}$ effector functions *in vitro*¹², without inducing death of these cells².

Earth science

The elusive coldfinger

George Helffrich

It was a penchant for canasta that led to the undoing of a world domination plan in an Ian Fleming novel. Fiction is simpler than reality, but simple ideas constitute powerful scientific tests. In the Earth sciences, a straightforward idea that explains some features of continents — their oceanic free-board, their low heat flow and their gravity signature — is that they are cooler and less dense than the mantle surrounding them. A lot of a continent — hundreds of kilometres — is below the surface, so a cool continent should lower the mantle's temperature and perhaps even induce convective down-

welling there. But according to Li *et al.* (page 160 of this issue¹), these tell-tale features of the continental coldfinger are absent, so it doesn't dominate the world at all — or at least not continental North America, where the work was carried out.

The story begins with the realization, in the mid-1970s, that continents have roots variably extending down to 300 km (ref. 2). Samples of rocks erupted from these thick continental areas indicate³ that at depths of 150 km they are as much as 200 °C cooler than the mantle that underlies the oceans. Moreover, the seismic velocities under conti-

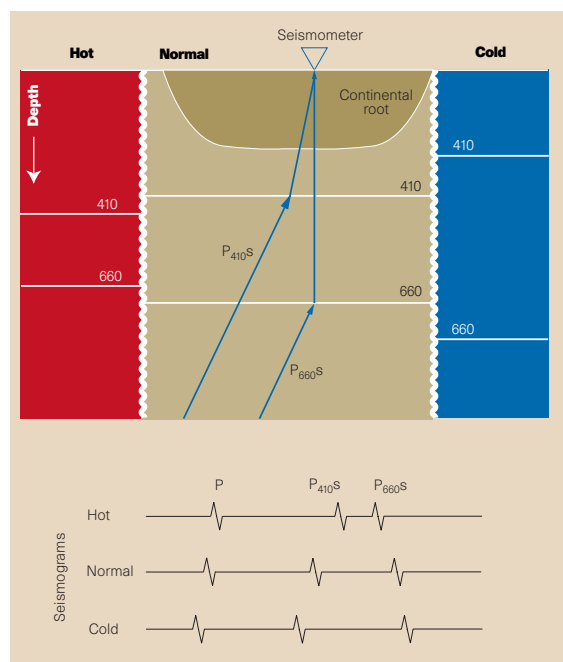


Figure 1 The approach used by Li *et al.*¹ to analyse mantle temperature in eastern North America.

Temperature changes in the mantle affect the depth of the seismic discontinuities nominally at 410 and 660 km. In a cold mantle, their spacing widens (right), and in a hot mantle it narrows (left). Seismic waves from distant earthquakes interact with the discontinuities, forming 'echoes' (P-to-S conversions at 410 and 660 km depth; P_{410S} and P_{660S}), whose time lags after the first-arriving P wave yield the conversion depth. This is one way to tell whether a continental root, which might extend down to 300 km (centre), cools the mantle.

nents are faster than the average elsewhere at the same depth, which is also compatible with cooler temperatures. So there are good reasons to expect a chilling subcontinental ambience.

Li *et al.*¹ take the temperature of the eastern North American mantle by measuring the seismic waves produced when the wavefields from distant earthquakes interact with the discontinuities at 410 and 660 km depth (Fig. 1). These changes correspond to mineralogical transitions in the mantle. Conveniently, in a cool mantle the level of the 410 transition rises and that of the 660 falls owing to the transitions' respective thermodynamic properties (and conversely in a warm mantle). This is a particular example of a broadly applicable technique by which the mantle's temperature and chemical structure can be gauged^{4–7}. The mantle transitions create 'echoes' of the main seismic waves, which, by their time lags after their predecessors, yield the depth to the transition. Thus, by seeing how these depths change where the continental root deepens, one probes its internal temperature.

During 1995 and 1996, an array of 18 portable seismic stations was deployed across the eastern United States, running from the northeastern seaboard into Missouri, to record distant earthquakes. The transect spans a region of deepening continental draft, and it is this region that Li *et al.* now report on. Astonishingly, the study finds no 410 deflection — at odds with the expected behaviour of a continental coldfinger. More paradoxically, on approaching the continental interior, the deeper discontinuity gets deeper still whereas the shallower one holds a steady level. This seems to preclude any large-scale, vertically coherent downwelling beneath eastern North America.

One escape from this conundrum is that

the downwelling could be time-dependent⁸ and that we are unlucky to be looking for it just now. Perhaps 50 million years of patient waiting will reveal one. A quicker verdict, barring mendacity on the Earth's part, could come from investigations of other continental interiors, which might catch downwelling elsewhere. But a worrying result from a sparser but broader-scale survey of North America using the same method reveals nothing more widely suggestive of cooler temperatures or downwelling⁹.

Alternatively, the methodology may suffer from unknown deficiencies of a frequency-dependent nature. Short-wavelength topography estimates on the corresponding discontinuities in the western Pacific differ by factors of up to four depending on the seismic frequencies used^{10,11}. Similar small-scale features may lurk unobserved under continents as well, and may be involved in mantle cooling. At present, we can only adopt a wait-and-see attitude akin to the Fleming character: once is happenstance; twice is coincidence; the third time is Earth-like behaviour. □

George Helffrich is in the Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK.
e-mail: george@olivine.geology.bristol.ac.uk

- Li, A., Fischer, K. M., Wyssession, M. E. & Clarke, T. J. *Nature* **395**, 160–163 (1998).
- Jordan, T. H. *Rev. Geophys.* **13**, 1–12 (1975).
- Boyd, F. R., Gurney, J. J. & Richardson, S. H. *Nature* **315**, 387–380 (1985).
- Revenaugh, J. & Jordan, T. H. *J. Geophys. Res.* **94**, 5787–5813 (1989).
- Vidale, J. E. & Benz, H. M. *Nature* **356**, 678–683 (1992).
- Helffrich, G. & Wood, B. *Geophys. J. Int.* **126**, F7–F12 (1996).
- Shen, Y., Solomon, S. C., Bjarnason, I. T. & Purdy, G. M. *Geophys. Res. Lett.* **23**, 3527–3530 (1996).
- Lenardic, A. & Kaula, W. M. *Nature* **378**, 709–711 (1995).
- Bostock, M. G. *J. Geophys. Res.* **101**, 8393–8402 (1996).
- Flanagan, M. P. & Shearer, P. M. *J. Geophys. Res.* **103**, 2673–2692 (1998).
- Castle, J. C. & Creager, K. C. *J. Geophys. Res.* **103**, 12511–12527 (1998).

Daedalus

Cruise micromissiles

The Crookes radiometer is a glass bulb containing a little 'mill' that spins in sunlight. Each vane of the mill is blackened on one face, which warms up compared with the other, reflective, face. Gas molecules hitting a vane edge flow back along it from the cold side to the hot one, and drive the vane by their reaction. This 'thermal transpiration' is most efficient when the mean free path of the gas molecules matches the vane thickness, which for vanes 0.1 mm thick implies a low pressure around 5 pascals.

Daedalus once devised a radiometer helicopter rotor, to fly in the low pressures at high altitude. He now points out that the mean free path in air at ground level is about 10 nm; so a radiometer could spin in ambient air if it had blades 10 nm thick. No macroscopic blade could be made so thin. But imagine, says Daedalus, a 10-nm wire, heated by a current or radiative absorption, and moving sideways through ordinary air. Its leading edge would be selectively cooled by the impacting air molecules, thus establishing a front-to-rear temperature difference. So those molecules would be impelled backwards by thermal transpiration, driving the wire forward and augmenting its speed. Given a starting push, the wire would continue to accelerate in that direction, like a ramjet.

So Daedalus's 'ram radiometer' is simply a fine grid of 10-nm wires in a duct about a millimetre across. It is the smallest possible aero-engine. Several such engines distributed about a tiny winged airframe will complete an 'artificial fly'. Modern microfabrication methods will shape the wire grids, form the tiny photovoltaic cells that power them, and lay down the steering electronics. This will turn different engines on and off in response to infrared digital codes beamed at the photocells.

Powered by daylight and controlled by an infrared unit rather like a TV-channel changer, the artificial fly will be an elegant and unobtrusive remote-sensing and transmitting gadget. Carrying a tiny microphone or a set of chemical sensors, it could snoop and soar, sniffing out leaks in a chemical plant, conversations in a crowd, or the microclimate in a forest. Daedalus likes the idea of forming a number of them into a personal defence squadron against real flies and mosquitoes. They could defeat the pests by ramming — insect wings are notoriously vulnerable to mechanical damage — or even, with sufficiently cunning control programs, by seduction.

David Jones

Audubon in action

Illustrators face a choice of truths: should they draw a creature in studio detail, or in its natural surroundings half hidden by shade? John James Audubon tried to capture the nature of birds in pictures backed by passionate prose.

Martin Kemp

Naturalism — showing it as it is — has been the declared aspiration of natural history illustrators since the Renaissance. But, faced with the limitations of static depictions in line and colour on a flat surface, compared with the dynamism of living forms seen in radiant light by a mobile observer, the illustrator has to make radical choices. For example, evoking the play of light across the feathers of a bird camouflaged in its natural environment excludes a careful description of its plumage and silhouette. One variety of naturalism necessarily prejudices another.

John James Audubon's majestic *The Birds of America*, issued in four volumes between 1827 and 1838, courtesy of a distinguished list of 161 subscribers in Europe and America, obviated one of the problems through sheer scale. His 'elephant' portfolio, a metre in height, allowed the life-size depiction of his 435 subjects, albeit with some markedly bent necks, legs and bodies for the bigger species. But all the other choices remained.

He operated almost exclusively with two registers of information. The first is the close-up observation of surface features, based on detailed scrutiny of dead birds, often shot for the purpose — "Alas, poor things," as even

such an enthusiastic 'sportsman' as Audubon lamented. His goal is measured description: "the compasses aided me... regulated and corrected each part". The second is the use of outline to capture the silhouette, motion and dynamic 'personality' of the bird in the field, as it might have been seen fleetingly or at a distance. Defending his use of *outré* poses, he stressed that "nothing can be more transient or varied than the position of birds".

Between these two registers, which play graphically to design and surface pattern, he eliminates much of the plastic description of the bird's three-dimensional form. His resources for the foreshortening of forms are very limited, and he exploits tonal modelling in light and shade only to a modest degree.

Audubon's emphasis upon surface design does not, as it might in other hands, lead to static or ornamental results. As a passionate field naturalist, more prone to romantic immersion than to systematic analysis, Audubon was concerned to capture the very nature of his subjects — the dynamics of lives dominated by feeding, survival, courtship and reproduction within the rich texture of their environments. The accompanying text volumes, his *Ornithological Biography*, combine sober description with expressions of pantheistic rapture characteristic of Romantic art and science.

His 'biography' of the mockingbird opens with a florid account of the luscious wonders of Louisiana, "where Nature seems... to have strewed with unsparing hand the diversified seeds from which have sprung all the beautiful and splendid forms which I should in vain attempt to describe". Fittingly for such an abode, the king of songsters conducts an ecstatic courtship, "pouring forth his melody, full of exultation at the conquest he has made".

In the more sober account that follows, Audubon outlines the natural history of the bird, noting that when marauding snakes attack a nest they are assaulted by "many other Mockingbirds from the vicinity". The tenor of his illustration — densely interweaving tree, plant, rattlesnake, birds and nest — is typical of the threat and malevolence that is present, overtly or covertly, in so many of his plates.

Appealing and decorative as his designs may be, the untamed world of action he depicts is closer to Darwin's than to that of the standard bird book. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.



Audubon's Mockingbird (*Mimus polyglottos*), from *The Birds of America*, Vol. I, 1931, plate XXI.

Perceptual bias for rising tones

An approaching sound source creates a pattern of rising intensity that can specify the arrival time of the source¹⁻³. Here we found that listeners reliably overestimated the change in level of rising level tones relative to equivalent falling level tones. In a natural environment this overestimation could provide a selective advantage, because rising intensity can signal movement of the source towards an organism. The bias was stronger at higher levels, suggesting that rising loudness is even more critical when a sound source is either close or loud. These results suggest a privileged status of dynamic rising loudness for harmonic tones and an asymmetry in the neural coding of harmonic dynamic intensity change.

We played synthetic vowel sounds that rose or fell equally in level to 12 subjects (Table 1). Listeners indicated change in loudness by positioning a cursor along an unmarked visual analogue scale, without considering either the direction of change or overall loudness. The left end of the scale was labelled "no change" and the right end "large change".

Rising level tones seemed to change more than falling level tones despite having the same actual change in level ($F_{(1,11)} = 32.74$, $P < 0.001$; Fig. 1a and Table 2), indicating that direction of change is an important (and previously unaddressed) factor in the perception of dynamic loudness change between two stimulus intensities. Furthermore, loud sounds seemed to change more than soft sounds despite having the same change in level ($F_{(3,33)} = 40.34$, $P < 0.001$). A significant interaction showed that the difference between rising and falling level sounds was greater for loud sounds than for soft sounds ($F_{(3,33)} = 12.81$, $P < 0.001$). None of these findings is predicted by traditional psychophysical laws derived using static stimuli, indicating that there are dramatic differences between static and dynamic loudness perception.

We then twice replicated the procedure with 12 new listeners in each experiment, using 1-kHz sinusoids in one experiment and white noise in the other. As was found with vowel sounds, rising level sinusoids were perceived to change in loudness more than falling ones ($F_{(1,11)} = 25.81$, $P < 0.001$; Fig. 1b), and loud sinusoids more than soft

Table 2 Means and standard deviations for each condition.

	60-75	75-60	65-80	Intensity change (dB)				
				80-65	70-85	85-70	75-90	90-75
Complex tone								
Mean	30.99	18.80	47.04	21.58	62.34	23.80	72.41	36.79
s.d.	16.80	13.72	12.78	11.26	14.51	10.81	13.34	17.67
Sinusoid								
Mean	39.46	28.43	55.68	39.78	70.09	39.48	79.90	51.16
s.d.	15.83	21.62	14.17	12.92	10.53	15.03	12.75	19.99
White noise								
Mean	23.93	31.36	38.38	36.94	52.91	49.10	61.28	51.53
s.d.	12.64	17.63	14.34	16.84	19.88	10.22	24.28	16.80

Visual analogue scale units.

ones ($F_{(3,33)} = 25.92$, $P < 0.001$). The difference between rising and falling level sinusoids was greater for loud sounds than for soft ones ($F_{(3,33)} = 6.40$, $P < 0.005$). Perceived change in loudness for rising and falling level white noise was not significantly different ($F_{(1,11)} = 0.08$; Fig. 1c). However, noise at high levels changed in loudness more than noise at low levels despite having the same actual change in level ($F_{(3,33)} = 27.21$, $P < 0.001$). There was also a small but significant interaction of change direction and intensity range ($F_{(3,33)} = 3.54$, $P < 0.05$).

These results suggest that there is an asymmetry in the neural coding of harmonic rising and falling intensity sweeps,

either in the response patterns of the peripheral auditory system, or in the summation of these patterns higher in the auditory pathway. The asymmetry occurred with harmonic sounds but not with broadband noise. This may be consistent with dynamic localization priorities in a natural environment. Harmonic sounds are produced by a wide variety of biological sources⁴, and harmonic structure can facilitate source segregation and identification^{5,6}. Approaching such a source, or anticipating its approach, can be a critically important environmental event. However, naturally occurring continuous broadband noise is less common, and can be the result of multiple sources such as crowd noise, or dispersed phenomena such as wind or rain. Dynamic localization of these types of sources may not be as important.

Behavioural research has shown that listeners can use dynamic intensity change to guide locomotion and to anticipate the arrival of a sound source^{7,8}. Physiological research has shown that more intense stimuli have a lower neural response latency and higher response amplitude^{9,10}. Taken together these findings suggest that dynamically changing intensity is an environmentally important cue, and that loud (or close) sounds have greater perceptual importance than soft sounds. Preferential responses to rising intensity harmonic sounds may provide an organism with an advantage in preparing for contact with a sound source, or an increased margin of safety on its approach. The increased perceptual disparity between rising and falling level tones found at higher intensities suggests that the louder the source, or the closer a source is to the organism, the more important rising intensity becomes. Although many stimulus parameters have yet to be investigated, these results suggest that rising intensity harmonic sounds have perceptual priority.

John G. Neuhoﬀ

Department of Psychology, Lafayette College,
Easton, Pennsylvania 18042-1781, USA
e-mail: neuhoffj@lafayette.edu

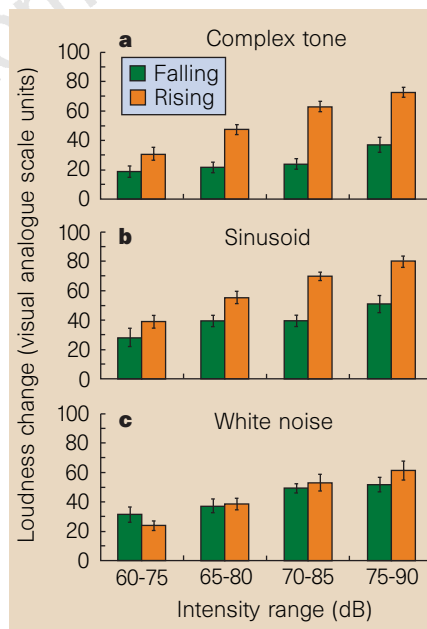


Figure 1 Perceived loudness change for each condition in each experiment. Stimuli were generated by a 16-bit sound card, fed into Sony MDR-v600 headphones, and had 10-ms onset and offset ramps. Sampling rate was 44.1 kHz for white noise and sinusoids, and 8 kHz for synthetic vowel tones. The vowel fundamental was 100 Hz with formants at 450, 1,450 and 2,450 Hz. Listeners received eight randomly ordered practice trials (one of each stimulus type) followed by each stimulus 10 times in random order. Means for each listener in each condition were used in the analysis.

Table 1 Conditions of intensity change.

Rising		Falling	
Onset	Offset	Onset	Offset
60	75	75	60
65	80	80	65
70	85	85	70
75	90	90	75

Intensity changed continuously (linearly in dB s⁻¹) between onset and offset over a stimulus duration of 1.8 s.

1. Shaw, B. K., McGowan, R. S. & Turvey, M. T. *Ecol. Psychol.* **3**, 253-261 (1991).

2. Lee, D. N., Vanderweel, F. R., Hitchcock, T., Matejowski, E. & Pettigrew, J. D. *J. Comp. Physiol. A* **171**, 563–571 (1992).
3. Neuhoﬀ, J. G. & McBeath, M. K. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 970–985 (1996).
4. Busnel, R. in *Animal Communication* (ed. Sebeok, T. A.) (Indiana Univ. Press, Bloomington, 1968).
5. Bregman, A. S. *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA, 1994).
6. Handel, S. in *Hearing: Handbook of Perception and Cognition* (ed. Moore, B. C. J.) (Academic, New York, 1995).
7. Ashmead, D. H., Davis, D. L. & Northington, A. *J. Exp. Psychol. Human Percept. Perform.* **21**, 256–269 (1995).
8. Rosenblum, L. D., Carello, C. & Pastore, R. E. *Perception* **16**, 175–186 (1987).
9. Sugg, M. J. & Polich, J. *Biol. Psychol.* **41**, 255–269 (1995).
10. Burkard, R. & Moss, C. F. *J. Acoust. Soc. Am.* **96**, 801–810 (1994).

p14^{ARF} links the tumour suppressors RB and p53

Most human cancers show perturbation of growth regulation mediated by the tumour-suppressor proteins retinoblastoma (RB) and p53 (ref. 1), indicating that loss of both pathways is necessary for tumour development. Loss of RB function leads to abnormal proliferation related to the deregulation of the E2F transcription factors, but also results in the activation of p53, which suppresses cell growth. Here we show that E2F-1 directly activates expression of the human tumour-suppressor protein p14^{ARF} (the mouse homologue is called p19^{ARF}), which binds to the MDM2–p53 complex and prevents p53 degradation^{2–5}. These results complete a pathway linking abnormal proliferative signals, such as loss of RB, with the activation of a p53 response, through E2F-1 and p14^{ARF}. They suggest that E2F-1, a protein inherently activated by cell-cycle progression, is part of a fail-safe mechanism to protect against aberrant cell growth.

We established inducible expression of E2F-1 in Saos-2 cells, a p53-null human osteosarcoma cell line that expresses low but detectable levels of endogenous p14^{ARF}. Activation of E2F-1 expression in these cells resulted in accelerated entry into DNA synthesis and apoptotic cell death (data not shown), as previously described following the transient expression of E2F-1 (ref. 6). Western blot analysis showed that the rapid increase in E2F-1 level following induction was closely followed by a dramatic increase in the level of p14^{ARF} (Fig. 1a). Such an increase in p14^{ARF} was not seen following the induction of two transcriptionally inactive E2F-1 mutants — E2F-1(1–374), a deletion mutant that lacks the transactivation domain, and E2F-1(132E), a point mutant that is unable to bind DNA⁷ (Fig. 1a). Both of these mutant E2F-1 proteins have lost the ability to enhance cell-cycle progression, although E2F-1(1–374) retains p53-independent apoptotic activity⁶. Levels of cyclin A, which is also the product of an E2F-regulated gene, remained unchanged

in these cycling cells following the induction of wild-type E2F-1. This finding is consistent with the observation that the greatest changes in expression of cyclin A would be observed in cells re-entering the cell cycle from quiescence. This indicates that the activation of p14^{ARF} expression by E2F-1 is not simply a consequence of cell-cycle progression. Expression of the two E2F-1 mutants, which inhibited cell-cycle progression, also led to the expected decrease in cyclin A expression.

The activity of the E2F-1 mutants suggests that activation of p14^{ARF} expression is dependent on the transcriptional activity of E2F-1. Analysis of the sequence of the region upstream of human exon 1β, which contains the initiation codon for p14^{ARF}, revealed the presence of a potential E2F-1 binding site, GCGGGAAA (see Supplementary information). We therefore isolated sequences encompassing the human exon 1β promoter and created a reporter plasmid in which the expression of the luciferase

gene was regulated by p14^{ARF} promoter sequences. Co-transfection experiments showed that, although wild-type E2F-1 efficiently activated transcription from this promoter, a transactivation-defective E2F-1 mutant had no activity (Fig. 1b). To confirm the transcriptional activation of the endogenous p14^{ARF} gene by E2F-1, we performed northern blot analyses of the E2F-1-inducible cells (Fig. 1c). As expected, E2F-1 transcription increased within 4 hours of induction, followed by a significant induction of the p14^{ARF} message. These results show that E2F-1 expression results directly in the transcriptional activation of p14^{ARF}, although we have also observed increased p14^{ARF} stability following E2F-1 activation (unpublished data).

We propose a model in which E2F-1, by activating p14^{ARF}, protects cells from oncogenic changes that result in abnormal proliferation (Fig. 1d). Our results provide a role for p14^{ARF}, which is not required for DNA damage-induced stabilization of

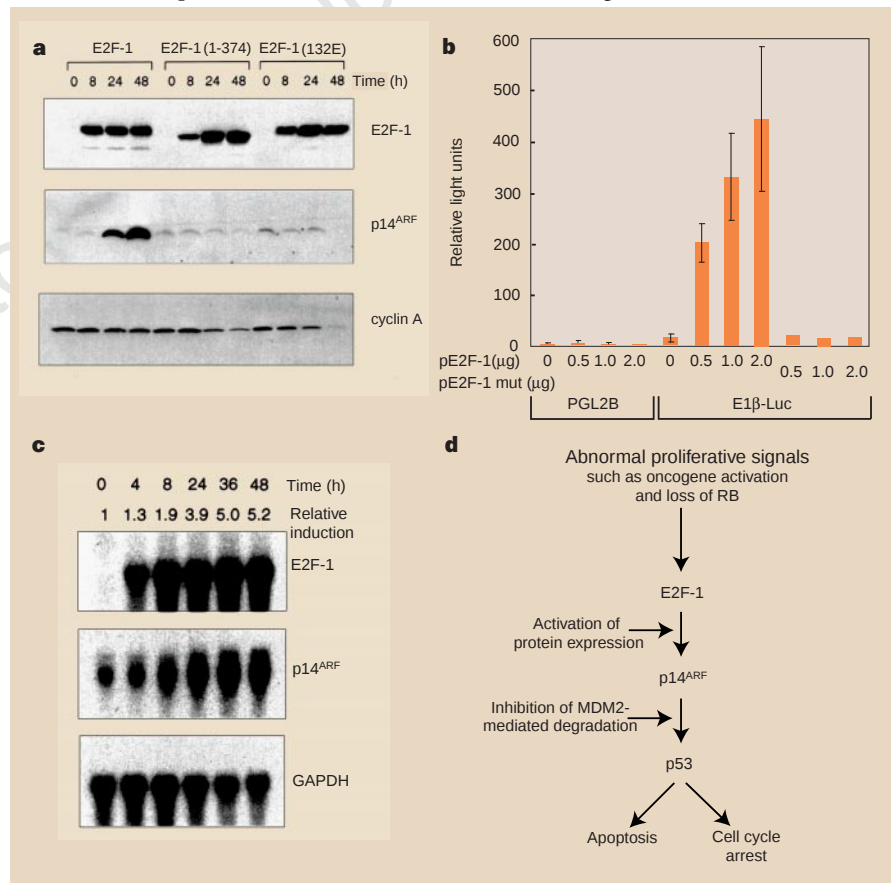


Figure 1 E2F-1 induced p14^{ARF} protein expression. **a**, Western blot analysis of Saos-2 cells inducibly expressing wild-type E2F-1, E2F-1(1–374) and E2F-1(132E). The expression of E2F-1, p14^{ARF} and cyclin A was monitored at the indicated time points following induction of E2F-1 by doxycycline. **b**, E2F-1 directly activates expression from the p14^{ARF} promoter. Activation of a reporter luciferase gene (E1b-Luc) under the control of the exon 1b promoter region following the expression of increasing amounts of wild-type, but not transcriptionally inactive, mutant E2F-1. E2F-1 does not activate transcription from a control reporter construct lacking the exon 1b sequences (PGL2B). **c**, E2F-1 induced p14^{ARF} mRNA expression. RNA was collected from the E2F-1-inducible cells at the indicated time points after the induction of E2F-1 by doxycycline, and northern blot analysis was carried out using probes to E2F-1 and p14^{ARF}. Equivalent loading was confirmed using a GAPDH probe. Quantification of the increase in p14^{ARF} expression following normalization for GAPDH is also shown. **d**, Model depicting the pathway linking RB and p53.

p53 (refs 4, 8). This link between E2F-1 function and p14^{ARF} expression provides an explanation for the ability of oncogenic stimuli that could deregulate E2F-1 activity, such as defects in the RB pathway or activation of oncogenes such as Ras, E1A or Myc, to activate p53 (refs 9–12). Thus, abnormal proliferation, resulting in deregulated E2F-1 activity, would induce p14^{ARF} and stabilize p53 (Fig. 1d). This would lead to cell-cycle arrest or apoptosis unless a second lesion occurred, such as a mutation in p14^{ARF} or p53 itself.

This model is supported by the observation that tumours containing p14^{ARF} mutations can tolerate the retention of wild-type p53 (refs 4, 8), strongly suggesting that loss of this pathway linking deregulated E2F-1 to the activation of p53 is an essential step in malignant progression.

Stewart Bates, Andrew C. Phillips

ABL Basic Research Program,
Chandler Street, Frederick,
Maryland 21702, USA

**Paula A. Clark, Francesca Stott,
Gordon Peters**

Imperial Cancer Research Fund Laboratories,
PO Box 123, 44 Lincoln's Inn Fields,
London WC2A 3PX, UK

Robert L. Ludwig, Karen H. Vousden

National Cancer Institute, Frederick Cancer
Research and Development Center,
West 7th Street, Frederick,
Maryland 21702-1201, USA
e-mail vousden@ncifcrf.gov

1. Palmero, I. & Peters, G. *Cancer Surv.* 27, 351–367 (1996).
2. Pomerantz, J. et al. *Cell* 92, 713–723 (1998).
3. Zhang, Y., Xiong, Y. & Yarbrough, W. G. *Cell* 92, 725–734 (1998).
4. Stott, F. et al. *EMBO J.* (in the press).
5. Kamijo, T. et al. *Proc. Natl Acad. Sci. USA* 95, 8292–8297 (1998).
6. Phillips, A. C., Bates, S., Ryan, K. M., Helin, K. & Vousden, K. H. *Genes Dev.* 11, 1853–1863 (1997).
7. Lukas, J., Petersen, B. O., Holm, K., Bartek, J. & Helin, K. *Mol. Cell Biol.* 16, 1047–1057 (1996).
8. Kamijo, T. et al. *Cell* 91, 649–659 (1997).
9. Morgenbesser, S. D., Williams, B. O., Jacks, T. & DePinho, R. A. *Nature* 371, 72–74 (1994).
10. Lowe, S. W. & Ruley, H. E. *Genes Dev.* 7, 535–545 (1993).
11. Hermeking, H. & Eick, D. *Science* 265, 2091–2092 (1994).
12. Kowalik, T. F., DeGregori, J., Leone, G., Jakoi, L. & Nevins, J. R. *Cell Growth Differ.* 9, 113–118 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

p19^{ARF} links the tumour suppressor p53 to Ras

Normal healthy cells possess safeguard mechanisms that sense oncogenic signals and trigger anti-tumorigenic responses that limit the proliferative potential of cells harbouring active oncogenes¹. In particular, expression of the Ras oncogene in normal primary cells causes a cell-cycle arrest that involves the activation of the tumour-

suppressor protein p53 (ref. 2). It has recently been reported that the tumour-suppressor p19^{ARF} can activate p53 (refs 3–5). Here we show that p19^{ARF} in the mouse (the human homologue is called p14^{ARF}) is essential for the activation of p53 in response to oncogenic Ras. These results, together with the finding that p19^{ARF} does not mediate the activation of p53 by DNA damage⁶, dissociate the activation of p53 into two pathways: one pathway is induced by DNA damage and is independent of p19^{ARF}, whereas the other pathway is induced by oncogenic Ras and is dependent on p19^{ARF}.

Under normal cellular conditions, p53 is maintained at relatively low levels because it has a short half-life. However, in the presence of DNA damage or oncogenic stresses, p53 accumulates rapidly through post-transcriptional mechanisms and gains full potential as a transcriptional activator⁷. The protein p19^{ARF}, one of the two tumour-suppressors encoded by the *INK4a-ARF* locus⁸, activates p53 both by neutralizing MDM2, which destabilizes p53, and by interacting directly with p53 (refs 3–5). DNA damage does not increase the level of p19^{ARF}, and the activation of p53 in response to DNA damage does not require p19^{ARF} (ref. 6). These observations led us to explore the possibility that p19^{ARF} is involved in the activation of p53 in response to oncogenic stimuli.

We have analysed the effect of an oncogenic form of the Ras protein (H-RasG12V, hereafter abbreviated as RasV12) on the steady-state levels of p19^{ARF} mRNA (Fig. 1a). The levels of p19^{ARF} mRNA were increased significantly (between five- and tenfold) by RasV12 (Fig. 1a). To test the role of p19^{ARF} on the accumulation of p53 triggered by RasV12, we measured endogenous p53 levels in primary mouse embryonic fibroblasts (MEFs) derived from wild-type and *ARF*-null embryos⁶ (provided by C. Sherr). As previously described², we found that the levels of p53 were increased significantly by RasV12 in wild-type MEFs (Fig. 1b, compare lanes 1 and 2). In contrast, p53 levels in *ARF*^{-/-} MEFs remained unchanged in the presence of RasV12 (Fig. 1b, compare lanes 3 and 4). These results show that p19^{ARF} is required for the accumulation of p53 in response to oncogenic Ras.

We have also determined the effect of RasV12 on the proliferation of wild-type and *ARF*^{-/-} MEFs (Fig. 1c). Wild-type cells infected with a retrovirus expressing RasV12 had significantly reduced proliferative activity, and most adopted a flat and extended morphology with cytoplasmic vacuolization (not shown), all features of senescence-like arrest². In contrast, whole cell populations of *ARF*^{-/-} MEFs expressing RasV12 proliferated at rates higher than

their vector-infected counterparts. In this respect, *ARF*^{-/-} cells expressing oncogenic RasV12 behave similarly to *p53*^{-/-} cells². These results extend the previously reported observation⁶ that monolayers of *ARF*^{-/-} MEFs transfected with oncogenic Ras develop foci of tumorigenic cells.

The induction of the tumour-suppressor protein p16^{INK4a} has also been implicated in the cell-cycle arrest elicited by oncogenic Ras². However, the complex genomic organization of the *INK4a-ARF* locus⁸ complicates the interpretation of some of the observations made with cells that lack p16 because of the elimination of exons 2 and 3, henceforth referred to as (*INK4a-ARF*)^{Δ2,3} cells (ref. 9). These cells retain exon 1β, which encodes for the active domain of p19^{ARF} (refs 4, 5, 10), and express an aberrant mRNA containing exon 1β (our unpublished observations). Therefore, it is possible that (*INK4a-ARF*)^{Δ2,3} cells produce

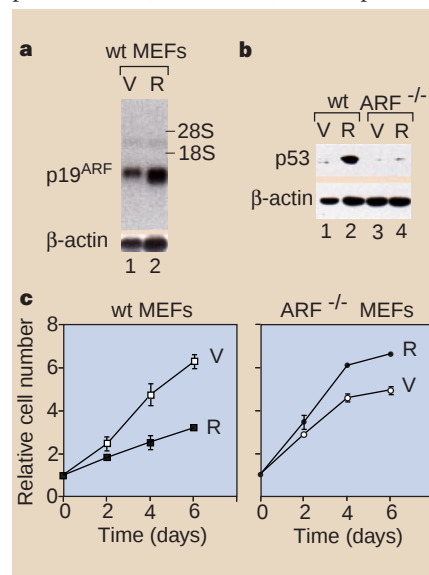


Figure 1 p19^{ARF} is essential for the p53-dependent arrest provoked by Ras. **a**, p19^{ARF} mRNA levels are increased in response to oncogenic Ras. Primary wild-type (wt) MEFs minimally passaged in culture were infected with replication-deficient retroviruses containing H-RasV12 (R) or the corresponding empty vector (V) (ref. 2). After selection of the infected cells, total RNA was extracted and probed with a labelled polymerase chain reaction fragment corresponding to exon 1β of the *INK4a-ARF* locus (top). To control for equal loading and transfer, blots were re-probed to detect actin (bottom). **b**, p53 accumulation requires p19^{ARF}. Primary MEFs of the indicated genotype were infected with empty vector (V) or H-RasV12 (R), as in **a**. Endogenous p53 levels were measured by immunoblotting with an anti-p53 antibody (top). To control for equal loading and transfer, blots were then re-probed to detect actin (bottom). **c**, *ARF*^{-/-} cells do not undergo proliferation arrest in response to oncogenic Ras. Growth curves of primary MEFs of the indicated genotype infected with empty vector (V) or H-RasV12 (R), as in **a**. The complete experiment was repeated three times, with similar results.

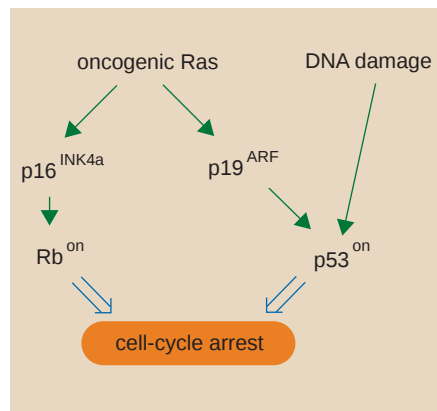


Figure 2 Anti-tumorigenic responses mediated by the products of the *INK4a-ARF* locus. Rb, retinoblastoma.

a peptide containing a functional p19^{ARF} domain. It has previously been reported that (*INK4a-ARF*)^{Δ2,3} cells do not arrest when oncogenic Ras is introduced but are still able to induce p53 (ref. 2). In the light of our present results, this suggests that (*INK4a-ARF*)^{Δ2,3} cells retain a functional, or partly functional, *ARF* gene. Taken together, these data indicate that oncogenic Ras elicits an anti-tumorigenic response mediated by the upregulation of both p19^{ARF} and p16^{INK4a}, which in turn activate the tumour suppressors p53 and retinoblastoma, respectively (Fig. 2).

Other oncogenes that are mechanistically unrelated to Ras have recently been reported to activate p53 in a p19^{ARF} dependent manner, resulting in a pro-apoptotic state^{11,12}. We speculate that p19^{ARF} senses unscheduled entry into the S phase of the cell cycle. Homozygous loss of the *INK4a-ARF* locus is common in human tumours. The results reported here, together with those of previous reports^{2,6,11,12}, indicate that the loss of p19^{ARF} and p16^{INK4a} renders cells unprotected against the action of oncogenes.

**Ignacio Palmero,
Cristina Pantoja,
Manuel Serrano**

Department of Immunology and Oncology,
Centro Nacional de Biotecnología,
CSIC, Campus de Cantoblanco,
Madrid E-28049,
Spain
e-mail: mserrano@cnb.uam.es

- Weinberg, R. A. *Cell* **88**, 573–575 (1997).
- Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. *Cell* **88**, 593–602 (1997).
- Pomerantz, J. et al. *Cell* **92**, 713–723 (1998).
- Zhang, Y., Xiong, Y. & Yarbrough, W. G. *Cell* **92**, 725–734 (1998).
- Kamijo, T. et al. *Proc. Natl Acad. Sci. USA* **95**, 8292–8297 (1998).
- Kamijo, T. et al. *Cell* **91**, 649–659 (1997).
- Levine, A. J. *Cell* **88**, 323–331 (1997).
- Haber, D. A. *Cell* **91**, 555–558 (1997).
- Serrano, M. et al. *Cell* **85**, 27–37 (1996).
- Quelle, D. E., Cheng, M., Ashmun, R. A. & Sherr, C. J. *Proc. Natl Acad. Sci. USA* **94**, 669–673 (1997).
- Zindy, F. et al. *Genes Dev.* **12**, 2424–2433 (1998).
- de Stanchina, E. et al. *Genes Dev.* **12**, 2434–2442 (1998).

Impacts on Earth in the Late Triassic

Spray *et al.*¹ postulate that five widely dispersed terrestrial impact structures with very similar geological age estimates (about 214 million years ago, in the Late Triassic epoch) are evidence of a multiple impact event. Most notably, the three largest impact structures, Saint Martin in western Canada (~40 km diameter), Manicouagan in eastern Canada (~100 km diameter), and Rochechouart in France (~25 km diameter), plot at virtually the same palaeolatitude in a continental reconstruction. Spray *et al.* suggest that this apparent crater chain was produced within hours as a series of coaxial projectiles collided in rapid succession with the rotating planet Earth, and drew analogies to the recent collision sequence of fragmented comet Shoemaker–Levy 9 with Jupiter.

However, published palaeomagnetic data for the Manicouagan^{2,3} and Rochechouart⁴ impact structures argue strongly against such a closely timed origin for these ancient events. This is because the characteristic remanent magnetizations of the melt rocks, including the most rapidly cooled glassy phases, indicate formation in a Late Triassic palaeomagnetic dipolar field of normal polarity at Manicouagan but of reverse polarity at Rochechouart.

These impact events must therefore have been separated temporally by at least the few thousand years⁵ it takes for a geomagnetic polarity reversal to take place, a process which in any case occurred relatively infrequently (at an average rate of about twice per million years^{6,7}) in the Late Triassic. Thus, although there is an interesting concentration of impact events in the Late Triassic, the opposite geomagnetic polarities recorded by the Manicouagan and Rochechouart melt rocks appear to preclude a synchronous multiple impact origin.

Dennis V. Kent

Lamont–Doherty Earth Observatory,
Columbia University,
Palisades, New York 10964, USA

- Spray, J. G., Kelley, S. P. & Rowley, D. B. *Nature* **392**, 171–173 (1998).
- Robertson, W. A. *Can. J. Earth Sci.* **4**, 641–649 (1967).
- Larochelle, A. & Currie, K. L. *J. Geophys. Res.* **72**, 4163–4169 (1967).
- Pohl, J. & Soffel, H. Z. *Geophysik* **37**, 857–866 (1971).
- Clement, B. M. & Kent, D. V. *Nature* **310**, 488–491 (1984).
- Gallet, Y., Besse, J., Krystyn, L., Theveniaut, H. & Marcoux, J. *Earth Planet. Sci. Lett.* **117**, 443–456 (1993).
- Kent, D. V., Olsen, P. E. & Witte, W. K. *J. Geophys. Res.* **100**, 14965–14998 (1995).

Spray replies — Kent raises an interesting point regarding the proposed late Triassic multiple-impact event¹. It might appear contradictory that the Manicouagan (~100

km diameter) and Rochechouart (~25 km diameter) impact structures possess normal and reversed geomagnetic reversals, respectively, if they were formed within hours of each other, as we suggested¹. However, palaeomagnetic fields are acquired when magnetic mineral phases pass through their Curie points (the temperatures at which iron minerals assume magnetic order and remain with their magnetic moments parallel to the Earth's magnetic field at that time). Critically, this does not necessarily coincide with the time of formation of the host rocks.

Rochechouart possesses a thin, sporadically developed impact melt layer (4 m thick at most). The palaeomagnetic data of Pohl and Soffel² were obtained from glass-bearing impact breccias and lithic breccias. The high-clast/low-melt content would have resulted in these rocks cooling below their Curie points (for example, approximately 580 °C for pure magnetite) within a short geological period, probably less than 100 years. This would be due to the cooling effect of the entrained cold clasts and the relatively rapid conductive and convective cooling of such a thin melt layer.

In contrast, Manicouagan, as a much larger impact structure, possesses an extensive and significantly thicker melt sheet (> 230 m and probably 500 m in total original thickness). Glass-rich breccias were not sampled as part of palaeomagnetic studies at Manicouagan^{4,5}, but crystalline melt sheet was sampled and many of the samples were medium-grained. It would have taken thousands, perhaps tens of thousands, of years for such a body of superheated melt to cool below the relevant Curie points. For example, Onorato *et al.*³ calculate that it would have required about 1,600 years for the centre of the melt sheet to reach its solidus (915 °C) and up to 10,000 years to cool to about 600 °C. This is for a body estimated at 200 m thick. We know now that 500 m is a more realistic thickness, so the estimates³ are on the low side.

Consequently, although both Manicouagan and Rochechouart could have been formed within hours of each other, the resulting impact-generated rocks would have reached their Curie points at times sufficiently different to allow for a natural geomagnetic reversal to have taken place, accounting for the different magnetic polarities of these impact structures.

John G. Spray

Department of Geology,
University of New Brunswick,
Fredericton, New Brunswick E3B 5A3, Canada

- Spray, J. G., Kelley, S. P. & Rowley, D. B. *Nature* **392**, 171–173 (1998).
- Pohl, J. & Soffel, H. Z. *Geophysik* **37**, 857–866 (1971).
- Onorato, P. I. K., Uhlmann, D. R. & Simonds, C. H. *J. Geophys. Res.* **83**, 2789–2798 (1978).
- Robertson, W. A. *Can. J. Earth Sci.* **4**, 641–649 (1967).
- Larochelle, A. & Currie, K. L. *H. J. Geophys. Res.* **72**, 4163–4169 (1967).

On-line journals and financial fire walls

The entire corpus of learned literature could soon be both on-line and free to the reader, with far-reaching consequences for academic life. What is standing in the way?

Stevan Harnad

The final state towards which the learned journal literature is evolving in the age of networked hypermedia is as inevitable as it is optimal: sooner or later, the entire corpus will be fully and freely accessible and navigable from the desk of any thinker in the world. Learned inquiry, always communal and cumulative, will not only be immeasurably better informed, new findings percolating through minds and media almost instantaneously, but it will also become incomparably more interactive.

When will it come to pass? It could already have happened; everything is in place, technologically speaking, and further resources are poised to follow suit whenever we are ready. But there is no second-guessing human nature; old habits find no end of rationales for perpetuating themselves and the status quo, even when, like an aged alpha wolf, they are far past the stage of being able to defend their status by anything but a menacing stare.

Unfounded fears

We will keep hearing solemn worries about (1) the preservability of texts in the new medium. In reality, however, nothing is simpler and more natural — particularly as more and more of our intellectual goods take flight from the paper flotilla into the digital skies, and our vested interests up there become ever more collective — than to arrange jointly for their continuous, systematic uploading and upgrading *pari passu* with ongoing developments in the medium. Bits are bits, and far more readily transferable than flecks of ink.

(2) The less-than-optimality of today's most advanced screen-reading technology for bath-, bed- and beach-based reading is hardly relevant to the desk-based searching, skimming, spot-checking, citation-tracing, and active cut-pasting and quote/commenting that is the mainstay of learned inquiry.

But the biggest brake on progress is still surely the reluctance of authors to entrust their work to a new, unproven medium in place of the one that has served them faithfully for centuries. Nor is their resistance based primarily on worries about preservability or readability. Authors are concerned about (3) quality control (will the new medium be as reliable and rigorous as paper?); (4) credit (will it bring recognition and advancement as paper did?); and (5) plagiarism (will it make work more vulnerable to theft?). The answers here are just as transparent. (3) Peer review, scholarship's classical mechanism of



MARK DOBSON

quality control, is medium-independent: it can be implemented in the air as readily as on land or sea. (4) Credit, likewise medium-independent, follows the best work and workers, it does not lead them (though authors and media can be like horses and water); kudos is ready to be assigned to skywriting as soon as authors are ready to redirect their pens heavenward. (5) Copyright, too, is medium-independent; and whatever increased power the Net may provide for stealing texts, it more than matches with its power for tracking them down.

Uploading the flotilla

Yet, despite these five *prima facie* obstacles, things may look to be moving along quickly enough: the Association of Research Libraries (ARL) *Directory of Electronic Journals, Newsletters and Academic Discussion Lists* (<http://www.arl.org:591>) listed over 3,400 electronic serials in 1997, twice as many as in 1996, 1,002 of them refereed (compared to 47 in 1996). This 2,000 per cent increase comes largely from the fact that more and more existing paper journals are now making electronic versions available for a fee: ARL listed 708, though this was already an underestimate, as last year's *Nature* review of new electronic journals (389, 137–138; 1997) had already counted over 2,000 such electronic *doppelgängers*.

But just a tenfold jump in the number of refereed journals on the Net from 1997–98, even if based more on second incarnations of

existing paper journals than on new launches, would go well over the 6,500 indexed by the Institute for Scientific Information and near the 14,000 indexed by Ulrich's *International Periodicals Directory*. That is indeed beginning to sound as if the core corpus may soon be in the sky for one and all.

For one and all? Recall that the optimal and inevitable was described as 'fully and freely accessible', whereas we are observing a proliferation of fee-based incarnations of paper journals on the Net, with financial fire walls separating them from one another as well as from the user. This issue of charging — and it is a crucial and controversial one — needs to be faced head-on.

The advocates of the status quo hold that things are progressing just as they ought. Both paper and on-line versions of journals are obviously still in demand, so let supply and demand orchestrate the transition and phase out paper if and when the market demands. There has been some consideration of pricing mechanisms other than subscription (site-licence and pay-per-view: S/SL/PPV), but the overall revenue per published page of journal has been held more or less fixed in the reckoning.

What about the cost? Here estimates have diverged: paper journal publishers, diversifying towards hybrid paper/on-line editions, estimate that costs per page increase (particularly as value is added to the on-line pages in the form of hyperlinks, multimedia, and so on), so there is certainly no basis for price

reduction. What about when demand shifts towards on-line only? It is a truism in publishing that expenses must be reckoned in terms of 'first-page' costs: it is all the work (writing, refereeing, editing, composition, mark-up) that goes into the first page that represents the lion's share of the expense. Printing and distributing multiple copies of that first page represents at the very most 30 per cent of the page cost, and if and when that 30 per cent saving can be made, publishers will be happy to pass it on to subscribers; for now, though, costs are up, not down.

The other side of the story comes from the minority of journals that have started as on-line only. There are as yet no exactly comparable pairs of paper-only and on-line-only journals, in terms of submission/acceptance rates, page counts, subject matter, readership, authorship, impact factor (citation ratio), and the all-important 'prestige' factor, but such comparisons as have been made suggest that on-line journals are managing for at least 70 per cent less, and that these economies could even be improved on as global archives come into the picture, scaling down the first-page costs to those of implementing (not performing) peer review (just as authors write gratis, referees referee gratis, and always have), plus editing/copy-editing and mark-up (much of that becoming increasingly off-loadable to authors too, as friendly HTML and eventually SGML authoring tools appear and pressure for their disciplined use by all authors grows).

Who pays the piper?

A great deal rides on this discrepancy in estimates of the true cost difference between paper and on-line-only pages. If the saving is really only 30 per cent or less, then S/SL/PPV is indeed the only way to recover costs, and free access, though still optimal, becomes unattainable instead of inevitable—absent a massive subsidy (from whom? to whom?).

But if the savings are the 70 per cent or more that the new e-only journal publishers are experiencing, then the optimal is not only inevitable, but within reach: currently, that 70 per cent is paid in large part by library subscriptions. Simple arithmetic shows that if the page charges for the remaining 30 per cent were paid in advance, at the author's end, out of (and for the sake of) those savings, not only would institutional libraries be much better off, but the world learned community — authors and readers alike — would be fast-forwarded straight into the optimal.

Unfortunately, the world learned community (as our earlier 'from whom?/to whom?' question suggests) is not a collective with any coherence or clout. Moreover, even if universities re-channelled all savings from library serials cancellations to faculty publication budgets, there would not be a balanced quid pro quo, because highly research-active universities (the net page-providers,

currently) would then face the biggest costs, whereas less research-active universities (the net page-consumers, currently) would get a free ride (although there might well be enough funds to keep it all aloft, because the research-active universities also tend to have the biggest serials collections; the 70 per cent savings still leave considerable room for manoeuvring; and, research productivity being the lifeblood of learned inquiry, and publication being its most measurable life-sign, other sources of research support, university as well as governmental, would find it natural to wrap into the funds for conducting the all-important research the relatively minor marginal cost of reporting it).

But, as long as restructuring depends on a collective — and an interdisciplinary, interinstitutional and international one at that — the optimal and inevitable may have a long wait. What is needed is some incentive to take matters into one's own hands as individual members of the learned community. There is a way, and it would allow individual scholars to have their cake and eat it too. The proposal is simple, and subversive. All authors should continue to entrust their work to the paper journals of their choice. But if, in addition, they were to publicly archive their pre-refereeing preprints and then their post-refereeing reprints on-line on their home servers, for free for all, then the de facto practices of the reader community would take care of the rest (irrespective of their reservations about bed/bath/beach reading); library serial cancellations, the collapse of the paper cardhouse, publisher perestroika, and a free for all, e-only serial corpus financed by author-end page charges would soon follow suit.

A centralized variant of this subversion scenario, xxx.lanl.gov, has already passed the point of no return in physics and some allied disciplines in the form of Paul Ginsparg's Physics Eprint Archive at Los Alamos National Laboratory, supported by the US National Science Foundation (NSF) and Department of Energy (DOE). As history will confirm, Ginsparg single-handedly set the world learned community on its inexorable course towards the optimal and the inevitable in August 1991. Hence 1998 e-

journals are not the right beans to count for e-future-casting: the number of new papers being deposited by authors at xxx (that subversive sapphire among the look-alike cyber-smut sites) is now 100 per weekday, with the number of 'hits' by readers now 65,000 daily and both figures still growing steadily. The real measures of the radical new way the world physics community is accessing its research literature are these, not the rate at which the fee-based on-line *doppelgängers* are being put on the market.

xxx marks the spot

In response to this reality, the forward-looking American Physical Society, publisher of some of the most prestigious physics journals in the world, has agreed to collaborate with xxx, which is already the de facto *locus classicus* for much of the physics literature. Manuscripts can be submitted to APS journals through xxx for refereeing; the final, refereed, edited version can then also appear through xxx, but with APS certification. These are the advantages of centralization. Backed by APS and NSF/DOE (and backed up at numerous mirror sites the world over), the preservation of this collective corpus is assured.

But the nature of cyberspace is such that there is always room for more. Mathematics, computer science and cognitive science (cogprints.soton.ac.uk) are already joining the disciplines served by xxx; there is no reason whatever why xxx should not go on to subsume (or subserve, rather) all the rest of the learned serial literature too; economies of scale suggest that this may even be the most efficient and direct path to the optimal; and a transition to page-charge-based cost-recovery and free distribution in place of S/SL/PPV will mean that journals — competing for authors' papers rather than readers' payments — are free from the need for fire walls to segregate papers from readers, and can instead collaborate in implementing the optimal and the inevitable for learned inquiry. □

Stevan Harnad is in the Department of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK. See www.cogsci.soton.ac.uk/~harnad for references and links connected with this article.

Journals index

● AIDS and Behavior	132	● International Journal of	
● Applied Geographic Studies	132	Molecular Medicine	133
● Boreal Environmental Research	130	● Journal of Industrial Ecology	131
● Child Psychology and Psychiatry Review	129	● Mangroves and Saltmarshes	131
● Combustion Theory and Modelling	135	● Microscopy and Microanalysis	134
● Current Opinion in Chemical Biology	133	● Multibody System Dynamics	136
● Extremophiles: Life Under		● Neuroscience-Net	129
Extreme Conditions	136	● Nitric Oxide: Biology and Chemistry	133
● Field Analytical Chemistry and		● Nonlinear Dynamics, Psychology,	
Technology	134	and Life Sciences	130
● International Journal of		● Technical Tips Online	135
Historical Archaeology	135	● Trends in Cognitive Sciences	129

Gathering the strands of thought

Trends in Cognitive Sciences

Editor Peter Collins

Elsevier. 12/yr. Europe £475, USA \$775 (institutional); Europe £85, USA £140 (personal)

Martha J. Farah

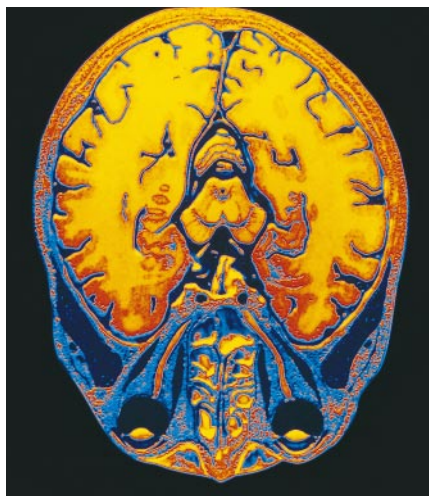
The late twentieth century has seen a number of fundamental transformations in the study of the mind. For decades, ending in the 1960s, behaviourism dominated psychology. Its basic tenet was that observable entities such as stimulus and response were the only legitimate subject matter of a scientific psychology, and that the mental states intervening between stimulus and response were beyond the scope of objective inquiry.

This view was challenged by the advent of cognitive psychology, sometimes dated to the 1967 publication of a book by that name by Ulric Neisser. The explicit goal of cognitive psychology was to characterize, in objective, mechanistic terms, the mental machinery underlying such cognitive processes as perception, attention, learning, memory, language and problem-solving. A key idea in cognitive psychology was the 'computer analogy', according to which the mind is to the brain as the software is to the hardware of a computer. When viewed this way, mental states no longer seemed the airy fairy stuff of behaviourist critiques, but information-processing states of a physical mechanism, and hence well within the grasp of science.

Within ten years, the phrase 'cognitive science' had entered our vocabularies, and marked another important turn in the history of the field. Cognitive science is an interdisciplinary endeavour encompassing cognitive psychology, computer science, neuroscience and linguistics. The relevance of computer science and neuroscience is obvious for a field defined by its mind-brain/software-hardware analogy. Linguistics also became involved, partly because language is one of the more spectacular forms of cognitive ability possessed by humans, and partly because some believe that natural language is a representational system that shares properties with mental representation.

If ever a field needed a journal devoted to reviewing current trends, it would be a new interdisciplinary field like cognitive science. Only the youngest practitioners have had the benefit of formal training in more than one or two of its component disciplines, the rest of us having retooled as best we could mid-career. And whatever one's training, it would be difficult to keep entirely up to speed with all four strands of the subject. *Trends in Cognitive Sciences* is therefore most welcome.

Each issue features state-of-the-art reviews of a variety of areas in cognitive sci-



ence, mostly by well-known authorities. The editors have managed to enforce an accessible style of writing in most of the articles, and the overall look of the journal is attractive and fun. Indeed, whereas all of my other journals are shelved in my office at the university, my copies of *Trends in Cognitive Sciences* are piled on my bookcase at home. As solid as their scientific content is, they seem more like pleasure than business. □

Martha J. Farah is in the Department of Psychology, University of Pennsylvania, Philadelphia, PA 19122, USA.

Brains on call

Neuroscience-Net

Editor John E. Johnson, Jr.

Scientific Design and Information. Online only, free.

<http://www.neuroscience.com>

Todd C. Holmes

Delightful as books and journals are, Internet-based publishing provides new capabilities for conveying scientific information, including rapid publication, search and link functions, running commentary on previously published articles, and vivid graphics. Theoretical scientists can set up web-based demos that allow interactive exploration by the reader, while scientists concerned with structure or dynamic processes can publish eye-catching movies. And for those of us who equate photocopying an article with reading it, web browsers offer the technological improvement of bookmarking.

Neuroscience-Net has made the brave step into electronic-only publishing with a journal for the neuroscience community. The journal subscription is free and is available to anyone with a computer and web access. It covers an admirably broad range of topics, divided by sections that include anatomy,

pharmacology, molecular biology, physiology, psychiatry and psychology, and theoretical neuroscience.

Its editorial process is streamlined by FTP manuscript submission and e-mail peer review. *Neuroscience-Net* guarantees publication within two weeks of acceptance, a remarkable turnaround by any standard. A check of the date of acceptance on several articles confirms rapid publication. I found the associated graphics informative and clear, but the hardware and choice of web browser of individual readers will determine the final graphical quality. Unfortunately, *Neuroscience-Net* lacks linkage with other websites. Referenced papers are not linked to other journals, and other sections have unrealized potential for valuable links — the Materials and Methods articles, for example, could have links to reagent suppliers and instrument manufacturers (such links could provide potential revenue for an electronic journal as well). The articles are of high quality, but the most recently published at the time of writing was August 25, 1997.

Neuroscience-Net provides a forum for rapid publication of neuroscience articles, but it does not yet take full advantage of the capabilities of Internet-based publishing. This journal merits reading by neuroscientists of all stripes. □

Todd C. Holmes is in the Department of Biology, New York University, New York, NY 10003, USA.

Seen and heard

Child Psychology and Psychiatry Review

Editors Linda Dowdney and Stephen Scott
Cambridge University Press. 4/yr. USA \$112, elsewhere £70 (institutional); USA \$56, elsewhere £35 (personal)

Peter Bryant

There are two different reasons why psychologists and psychiatrists are interested in children. One, an academic reason, is their interest in the origins of people's behaviour. We are bound to be in a better position to understand any aspect of behaviour if we know how it started, and most of people's behaviour (their solutions to intellectual problems, their prejudices, their moral principles) can be traced back to childhood. The second reason is a practical one. Many children need help from experienced clinicians, and they cannot afford to wait until the problems that afflict them — for example, autism, abuse, dyslexia and school phobia — have been thoroughly analysed and understood by psychologists and psychiatrists.

Of course, there is a certain tension

between these two motivations, but on the whole it has been a creative tension. Academic psychologists have certainly benefited from the insights and experiences of people with clinical experience.

In this country, the enterprising and highly successful *Journal of Child Psychology and Psychiatry* has played an important part in making this contact between academic researchers and practitioners possible. For many years this journal has published scientific papers of a high standard, which usually interest people with clinical interests as well as academic readers.

Now the association that runs this journal has started a new enterprise: the *Child Psychology and Psychiatry Review*. Its purpose is to keep practitioners informed about the latest developments in research and in practice, and it sets out to do so in several ways. Each issue contains commissioned articles or sets of commissioned articles that provide up-to-date and highly readable reviews of a particular topic (for example, dyslexia, behavioural genetics). It also contains submitted articles and a journal monitor section with brief summaries of papers in other journals that are relevant to particular childhood problems, such as child abuse and educational difficulties. Most issues also have a section on legal matters. There is an occasional personal profile, each with an interview of a well-known child psychologist or psychiatrist. The new journal also contains a fairly lively correspondence section.

The articles are a mix of down-to-earth accounts of clinical practice and of up-to-date descriptions of theories on this or that aspect of childhood. They are all written clearly and with an attractive simplicity. The other sections of the journal are also interesting and informative, except for the personal profile section which has a parochial air and does not seem to me to add any stimulating new information. Apart from this, the editors of this new journal are providing busy practitioners with an entertaining but intellectually respectable account of current research. □

Peter Bryant is in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

The chaotic psyche

Nonlinear Dynamics, Psychology, and Life Sciences

Editor Stephen J. Guastello
Human Sciences Press. 4/yr. USA \$110,
elsewhere \$130 (institutional); USA \$35,
elsewhere \$41 (personal)

H. Eugene Stanley

The emerging field of nonlinear science has led to many advances and to many new journals — *Fractals*, *Nonlinearity* and *Physica D* to name just three. Recently, the concepts

and methods of nonlinear science have also been finding interesting applications in the social sciences. Witness, for example, the popularity of applying methods of statistical physics to problems of economics and finance, an endeavour that has spawned many interesting ideas and topical conferences in Budapest, Palermo and Dublin under the rubric of 'econophysics'.

Other topics in the social sciences that appear to be amenable to the concepts and methods of statistical physics include problems associated with urban growth, where one quantitatively analyses the locations of homes, or foraging phenomena, where one asks for the most efficient strategy to find an object (food or lost car keys).

In all this work, there is no attempt to model the free will of the subject that determines the outcome, be they a Wall Street trader, a citizen building a home, or the foraging birds and bees. Presumably individuals do not exercise their free will, but learn efficient strategies and implement them.

In the fields of psychology and psychoanalysis it is hard to deny the role of free will, hence a journal devoted to the applications of concepts of nonlinear science to these disciplines would seem most interesting. The papers collected in the first issues span a range of problems of general interest in psychology, and bring to bear on them concepts such as catastrophe theory and fractal theory. Out of these marriages arise children with unfamiliar names such as 'collective intelligence' and 'stochastic determinism' (an oxymoron at first sight).

Will this moderately priced journal fulfil the need for a genuinely interdisciplinary journal on nonlinear science applied to the social sciences? It seems doubtful. The editorial board contains no physical scientists or practitioners of statistical mechanics, the discipline most active in nonlinear science. Further, there is a strong emphasis on psychology rather than on a wider range of social sciences. Moreover, if the initial issues are any guide, there is a tendency in many of the papers to simply 'talk' rather than report new experiments, simulations or scientific

models. There are even a few misunderstandings of some of the principles of nonlinear science.

These limitations might be overcome, and the journal fulfil its potential, if it were to broaden its horizons and undertake rigorous refereeing procedures.

H. E. Stanley is at the Center for Polymer Science, Boston University, 147 Bay State Road, Boston, MA 02215, USA.

No waste of trees

Boreal Environmental Research

Editors-in-chief Hannu Lehtonen and

Sylvain Joffre

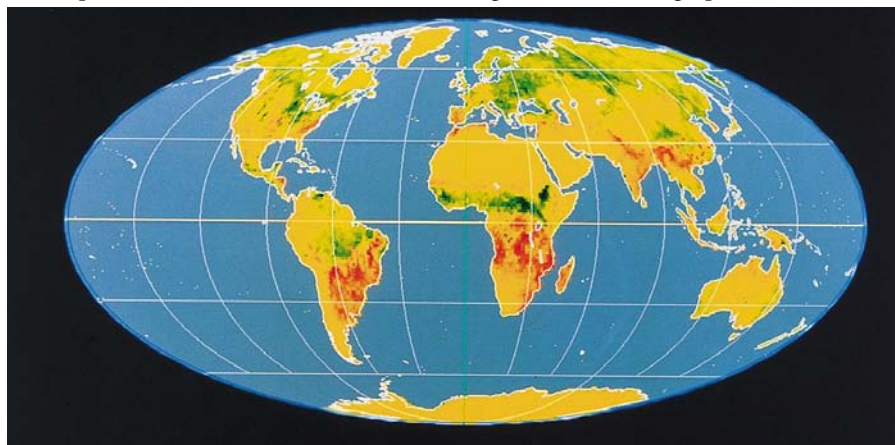
Finnish Zoological and Botanical Publishing Board. 4/yr. FM500, \$100 (institutional); FM200, \$40 (personal)

David W. Schindler

No single ecological type, or 'biome', covers more first-world countries than the Boreal, with forests that extend over most northerly latitudes of Eurasia and North America, including Scandinavia, the former Soviet Union, Canada, and the northern USA. For the past 30 years, residents and environmental organizations of these countries have expressed their displeasure at the rapid exploitation of tropical forests, not noticing that boreal ecosystems under their very noses were quietly being sacked at an equally astounding rate. The boreal zone also contains more than 90 per cent of the world's lakes. Its vast wetlands are pivotal ecosystems in greenhouse gas balances.

In the past, research on boreal lands and waters has been diffused through a number of more general ecological publications, making it difficult to find. A journal that advertises itself as devoted to boreal science in the broadest sense, emphasizing holistic approaches to boreal problems by experimental, monitoring, theoretical and modelling approaches, is therefore long overdue.

Volumes one and two of *Boreal Environmental Research*, from 1996 and 1997, make a good start in living up to that advertise-



The fixing of CO₂ (green) and its release through plant decomposition (red) in a typical June.

ment. Research on lakes, streams, wetlands, forests and the Baltic Sea is included. While most of the papers are by Finnish scientists, there are a fair number from Norway and Sweden and a sprinkling from Canada, China and other countries. The quality of research papers is comparable with those published in well-known ecological periodicals from other countries. I am confident that as the journal becomes more widely known, the authorship will become more broadly international. That being said, a journal that makes the excellent, but often quite inaccessible, boreal research in Finland easily available internationally is important in its own right.

The journal is remarkably well turned out for a new, modestly priced periodical. Figures and tables are clear and well reproduced, and the layout is very attractive. In contrast to many foreign-produced publications in English, the writing and editing is nearly flawless ... in fact, better than in most English language publications from North America.

My sole criticism is that several of the papers do not reference key relevant research papers from North America. This would easily be corrected by using referees from more countries, which will no doubt happen as the journal attracts a more international authorship. Publication is very rapid, on average about four months after acceptance for the volumes under review.

Perhaps my highest recommendation is that I have recommended that our library subscribe, despite declining budgets for journals. It is indispensable for those working in the Boreal, and very useful for those engaged in ecological research in other ecosystems. □

David W. Schindler is in the Department of Biological Sciences, University of Alberta, Edmonton T6G 2E9, Canada.

Common ground among quagmires

Journal of Industrial Ecology

Editors David Allen and John Ehrenfeld
MIT Press. 4/yr. \$95 (institutional); \$40 (personal)

Gilbert S. Hedstrom

The business world of the twenty-first century is likely to be shaped by now-familiar forces, including rapid technology development, greater openness and vitality in political systems, and rampant globalization. At the same time, a second set of forces has the power to strongly affect the direction it takes. These 'sustainability drivers', fuelled by increasing world population and declining living systems, may be the least appreciated, least understood set of threats and opportu-

nities facing business.

Industrial ecology is a discipline that provides for exploring this territory and exploiting this new set of opportunities. While the term 'industrial ecology' has been around for over ten years, this is essentially a new field. And like any new field, it has the potential for either bringing clarity and focus to significant issues or becoming mired in jargon and academic debate. For industrial ecology to be a useful part of sustainability solutions, the scientific, government and business communities will need to speak the same language, based not on self-interest and emotion but rather on common interests and logic.

The *Journal of Industrial Ecology* takes a major step in bringing clarity to a potentially very confusing topic. The quarterly journal, launched in the winter of 1997, offers a thoughtful and comprehensive, yet accessible venue for exploring the potential role of industry and government in reducing the environmental burdens throughout the product life cycle, from extraction of raw materials, to the production of goods, to the use of those goods and to the management of the resulting wastes.

Several factors distinguish this journal. It is timely, for to think about issues of sustainability as the next century dawns is both prudent and ethical. Its global perspective brings fresh and balanced voices to a sometimes parochial discussion. Finally, it is both very readable and very well edited, with the articles fitting together more like a puzzle than a melange of individual contributions.

Not to be read only once or in one sitting, this journal is an important reference for those interested in how business and government can maintain the benefits of the first industrial revolution while dramatically reducing the burden those benefits place on the underlying living systems that need to be sustained during the next industrial revolution. □

Gilbert S. Hedstrom is at Arthur D. Little, Inc., Acorn Park, Cambridge, Massachusetts 02140, USA.

Imperilled wetlands

Mangroves and Saltmarshes

Chief Editors Eric Wolanski and Charles S. Hopkinson Jr.
SPB Academic Publishing. 4/yr. Print only.
\$152 (institutional); \$80 (personal)

Thomas J. Smith III

Mangrove forests dominate the intertidal zone of the world's tropical coastlines, yielding to saltmarshes in the temperate and arctic zones. Unfortunately, more than 80 per cent of the world's population also lives in the coastal zone. Saltmarshes and mangroves face tremendous human pressure from



Coastal mangrove swamp at low tide.

development, despite their known ecological services to man. These services include: supporting productive fisheries; preventing shoreline erosion, especially from storms; and the uptake and transformation of nutrients. In some developing countries in the tropics, mangrove forest is being lost at a rate of more than 10 per cent a year. A journal devoted to the scientific understanding and enlightened management of these two important coastal ecosystems is therefore quite justifiable.

Since its launch in December 1996, *Mangroves and Saltmarshes* has published four issues, each with six research articles. Papers address both pure and applied research topics, but notes, comments, rebuttals, book reviews and so on are not included. The layout and graphical presentation are very good, and the turnaround time for submissions is quick. Based on the dates presented in the last two issues, the average paper takes four to five months from submission to acceptance.

Unfortunately, the quality of the papers in the issues I reviewed was highly variable; ranging from horrid (presumably on the principle that if a paper has been rejected everywhere else, send it to a new journal) to truly excellent. The quality is not related to the country of origin; indeed, the worst paper, in my opinion, was from a member of the editorial board.

A strength of the journal is that the editorial board is endeavouring to assist authors from less developed nations to present their results to the worldwide scientific community.

Many paradigms concerning mangrove forests were originally developed from studies in Florida, but these paradigms do not necessarily hold for mangroves worldwide. So the presentation of research conducted in the vast expanses of mangroves in less developed areas is needed. Authors from the developed world need to see these studies

and examples from elsewhere to help us take a less blinkered view.

Because mangrove and saltmarsh ecosystems are both wetlands, and found in estuaries, *Mangroves and Saltmarshes* may have difficulty finding a niche. Numerous established journals cover similar themes, including: *Estuaries*; *Wetlands*; *Aquatic Botany*; *Estuarine, Coastal and Shelf Science*; and *Coastal Zone Management Journal*.

I hope *Mangroves and Saltmarshes* succeeds. I believe it can fill an important role for the research, and especially, resource management communities in developing and developed countries alike. □

Thomas J. Smith III is at the US Geological Survey, Biological Resources Division, OE Building, Room 148, Florida International University, Miami, FL 33199, USA.

Human interest stories

Applied Geographic Studies

Editor Milton E. Harvey

John Wiley and Sons. 4/yr. USA \$155, elsewhere \$179 (institutional); USA \$70, elsewhere \$94 (personal)

Michael Batty

Geography emerged as an academic discipline at the height of imperialist expansion in the late nineteenth century, but by the 1960s it had lost its taste for describing regions and landscapes in meticulous detail in favour of more systematic inquiry. It divided into human geography, which drew on theory and method in the social sciences, particularly economics, and physical geography, which oriented itself towards the earth sciences.

The subject has softened its scientific stance since then, moving away from a search for strong theory, becoming more pragmatic

and, in the process, taking on a multitude of applications.

Applied Geographic Studies was launched last year in response to such developments. Building on a series of successful conferences that drew together a diversity of geographic applications, its content reflects contemporary human issues that have important geographical dimensions. Environmental problems, such as the spatial distribution of pollutants and their implications for disease and health care, are of concern; also relevant to this focus is the use of geographical models for marketing and retailing, for problems posed by the impact of ageing, and for the provision of housing. The journal covers population studies involving urbanization, migration and the current problems of global geopolitics, although so far there is little discussion about the traditional areas of urban planning and transportation.

Starting a journal such as this is a hazardous undertaking, because a competitor (*Applied Geography*) has existed for a dozen or more years, and several others specialize in the subject. But the editor has assembled an impressive editorial board of 32 well-known academic geographers who are involved in applications, and nearly half of these have been enticed to write in the four issues that make up the first volume.

The result is a useful and balanced collection of competent articles that reflect the intended range. The articles are quiet and unpretentious as is the style and presentation of the journal. If the editor and board are able to keep up this momentum, then *Applied Geographic Studies* will soon become an established outlet for geographic applications of quality. □

Michael Batty is at the Centre for Advanced Spatial Analysis (CASA), University College London, 1-19 Torrington Place, London WC1E 6BT, UK.

Civil defence in the war against AIDS

AIDS and Behavior

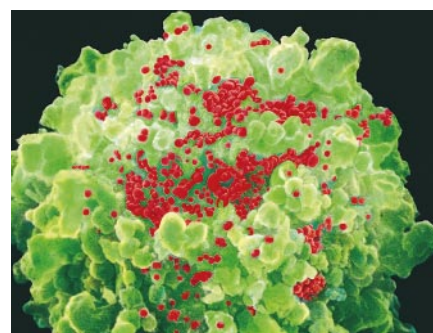
Editor Thomas L. Patterson

Plenum. 4/yr. USA \$150, elsewhere \$175 (institutional); USA \$50, elsewhere \$59 (personal)

Ralph J. DiClemente

While there have been great advances in the understanding and treatment of HIV, there is, at present, no cure. And, while promising candidate vaccines are being readied for testing, an effective, widely available vaccine is not anticipated in the foreseeable future. Thus, primary prevention of HIV transmission is of paramount importance in curtailing the pandemic.

The behavioural and social sciences have played and will continue to play a critical role



HIV particles budding from an infected T-cell.

in understanding the factors associated with HIV-related risk behaviours, both sexual and drug-related, and in the development of interventions designed to reduce such behaviours among HIV seronegative as well as HIV seropositive individuals.

The role of these sciences, however, is not limited to changing behaviour that might lead to infection. They also play a key role in understanding factors associated with medication adherence, with understanding an individual's coping response, with identifying the neuropsychological impact of HIV infection on cognitive functioning, and with examining the emotional effects of HIV such as depression. Such understanding can help in the development of interventions which will improve people's lives in all these areas.

HIV-related behavioural and social research was often published in broad-based public health journals or more clinically focused HIV publications, but the growth in such research makes *AIDS and Behavior* a welcome new vehicle for its dissemination.

The journal has an outstanding editorial board, a veritable Who's Who of the field. Many of the articles are empirically based rather than theoretical reviews or commentaries, but non-empirical articles are also welcomed. Emerging areas of research are more fully explored by groups of five to seven thematically linked papers co-ordinated by a guest editor, usually an expert in that particular area.

Topics span the gamut from primary to secondary prevention, adjustment to living with HIV, and the influence of a host of psychosocial factors on risk behaviour and adherence. The research is of high quality, perhaps reflecting the fact that many of the contributing authors are inveterate investigators. The journal is attractive and well formatted, with clear text, tables and figures. Articles are usually from eight to ten pages long, although longer articles are also published; turnaround time is satisfactory.

Anyone interested in this area of HIV research will be well served by examining *AIDS and Behavior*. □

Ralph J. DiClemente is in the Department of Behavioral Sciences and Health Education, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA.



Night lights: cities, burning vegetation (purple), oil flares (red) and the aurora borealis (light blue).

Rapid response

International Journal of Molecular Medicine

Editor and publisher Demetrios A.

Spandidos

12/yr. Europe \$600 (institutional), \$300 (personal); elsewhere \$650 (institutional), \$350 (personal)

Jeffrey M. Leiden

Basic scientists, physicians and venture capitalists all agree that twenty-first century medicine will be based increasingly on our understanding of the molecular genetic causes of human disease. The scheduled completion of the human genome project in 2005, along with advances in human genetics, immunology, molecular pathogenesis, structural biology and gene therapy are creating an interdisciplinary field of molecular medicine that promises to revolutionize the ways that we investigate, diagnose, and treat many acquired and inherited diseases.

We are already beginning to identify the genes involved in prevalent diseases such as hypertension, diabetes and cancer. The discovery of these genes and their functions is providing important pathophysiologic insights, and molecular probes can accurately assess a patient's susceptibility for disease long before they become symptomatic. Of equal importance, access to these genes and their protein products provides new targets for rational drug design, drug screening and somatic gene therapy.

This revolution in molecular medicine has created a vacuum in the biomedical literature. And, as is usually the case, medical publishers have rushed in to fill the void. New molecular medicine journals, including *Nature Medicine*, *Molecular Medicine*, *Human Molecular Genetics* and *Molecular Medicine Today*, publish original research articles and reviews in this area. Meanwhile, established journals such as the *Journal of Clinical Investigation* are increasingly filled with molecular medicine manuscripts. Given this cornucopia of literature, one could justifiably ask if we need yet another journal on molecular medicine. Only time will tell, but a review of the first five issues of the *International Journal of Molecular Medicine* suggests that the answer may be yes.

This journal's goal is the rapid publication of original work and reviews concerning the molecular mechanisms of human disease. Of necessity, this is a wide purview, encompassing the fields of human genetics, immunology, virology, pharmacology, pathology and gene therapy among others. Manuscripts can be submitted directly for peer review or can be contributed by members of the editorial board, who are experts in fields related to molecular medicine. The journal includes basic papers dealing with

the molecular pathways involved in pathological cell function, and more clinical papers describing novel pharmacologic and genetic therapies. Of equal importance, each issue contains three to five review articles that summarize a variety of interesting topics at the interface of molecular biology and clinical medicine. In general the articles are well written and the illustrations and graphics are first rate.

Compared with many other journals on the subject, the time from submission to acceptance is remarkably short, averaging less than one month in the issues under review. One of the most interesting aspects is the preponderance of articles submitted from outside the United States. This appears to reflect the international flavour of the editorial board and the increasing interest in molecular medicine in Europe and Asia. Some articles, however, could benefit from more careful grammatical editing. □

Jeffrey M. Leiden is in the Chief Section of Cardiology, University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois 60637, USA.

A fertile and dynamic sea

Current Opinion in Chemical Biology

Editors Donald Hilvert and Steven Ley
Current Biology. 6/yr. Print, £125, \$220 (personal); £60, \$120 (student). On-line £137, \$242 (personal). Print plus on-line, £150, \$264 (personal)

Brent L. Iverson

Chemical tools are increasingly being applied to, and chemists are increasingly being inspired by, the molecules of living systems as well as the living systems themselves. As scientists attempt to ride these rapidly advancing currents of progress, new paradigms quickly arise and become part of the fertile and dynamic sea of ideas and experimental results that make up chemical biology.

Current Opinion in Chemical Biology, first published in June 1997, covers timely and important topics of bioorganic as well as bioinorganic chemistry, along with relevant advances in enzymology and medicinal chemistry. Editors Donald Hilvert and Steven Ley have enlisted an impressive group of world leaders in the field to oversee the writing of short reviews by experts on recent developments in the most active areas of chemical biology research. Journal issues this year are devoted to interaction, assembly and processing (February); biocatalysis and biotransformation (February); bioinorganic chemistry (April); combinatorial chemistry (June); next generation therapeutics (August); analytical techniques (October); mechanisms (October); model systems (December); and biopolymers (December). The

review articles are of the highest scientific calibre and are focused on recent literature as viewed through the eyes of those who know and understand that literature best. The reviews are generally eight to ten pages in length, long enough to explore thoroughly all essential ideas, but not so long that the reader is set adrift in details.

Current Opinion in Chemical Biology also has 'paper alert' and 'web alert' sections to help readers negotiate the large number of new currents of chemical biology literature, both published and electronic. In these 'alert' sections, scientists representing the journal's key topic areas highlight important recent papers or outstanding websites, along with a concise summary of key results and their significance.

To ride new waves in chemical biology, scientists must know how recent scientific advances are moving around them. A subscription to *Current Opinion in Chemical Biology* will provide excellent guidance to scientists on how best to harness the onrush of new advances and trends, thereby preventing a scientific 'wipe out' caused by being overtaken from an important new direction by something they did not see coming. In scientific research, no good-looking lifeguard comes to the rescue. □

Brent L. Iverson is in the Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712, USA.

NO particular place to go

Nitric Oxide: Biology and Chemistry

Editor-in-chief Louis J. Ignarro
Academic Press. 6/yr. North America \$220, elsewhere \$262 (institutional); North America \$110, elsewhere \$132 (personal)

John Garthwaite

There is no doubting that the identification of nitric oxide (NO) as a biological signalling molecule has had an extraordinary impact. It all started about ten years ago when



Salvador Moncada: a pioneer in NO research.

researchers in three previously disparate areas of research (vascular biology, neuroscience and immunology) all stumbled on the same molecule.

Within a short period of time, it was realized that NO was involved in the regulation of blood vessels, in communication in the brain, and in immunological defence against invading organisms. Now, NO impinges on almost all areas of biology. Not only is it important physiologically as a messenger that conveys signals from one cell to another in many different tissues but, through its toxic effects, it is also a likely player in many pathological processes.

So it was decided that we need to have a journal devoted specifically to the molecule, *Nitric Oxide: Biology and Chemistry*. But do we? It could also be said that, because of its importance to so many areas of biology, a journal solely concerned with NO is a non-starter: NO is really only interesting in context. After all, we don't have, or need, journals specifically about glutamate or acetylcholine or inositol trisphosphate.

Viewed from this perspective, there is no shortage of journals in which to publish and read about NO research. In fact, almost any one will do.

Nitric Oxide: Biology and Chemistry is a spin-off from *Archives of Biochemistry and Biophysics* and is also referred to as 'Part B' of that journal. The first issue appeared in February 1997 and, so far, only eight issues have been published (six of which were from last year) so it is possibly a little early to expect the journal to have found its flavour.

The editors would like it to be a repository of papers about all aspects of the biology and chemistry of NO, but this is probably unrealistic. To date, there have been a few reviews, and a modest number (up to ten) of research papers per issue (except the latest: volume two, number two, which is a book of abstracts). Some of the better papers have come from the editorial advisory board's research groups. These aside, the journal appears to be falling into its default guise, which is *Archives of Biochemistry and Biophysics Part B*. The parent journal has long had a penchant for chemically oriented research into reactive free radicals, and NO does have this chemical characteristic (although it is not very reactive).

Of course it is important that the chemistry of NO and related species, such as the much more reactive peroxyxynitrite anion, does become better understood. Indeed, it is just this type of research that one imagines could find a home in a journal devoted exclusively to NO. The problem is that, when the chemistry is not constrained within a biological framework, you can get free radicals to do almost anything. □

John Garthwaite is at the Wolfson Institute for Biomedical Research, University College London, 140 Tottenham Court Road, London W1P 9LN, UK.

Venturing outside the laboratory

Field Analytical Chemistry and Technology

Editor Henk L. C. Meuzelaar
Wiley. 6/yr. USA \$195, elsewhere \$231

Patrick MacCarthy

As analytical chemistry has evolved, it has focused primarily on the determination of substances within the controlled confines of a fixed laboratory. While there has always been an interest in conducting analyses where a sample is actually found, such on-site measurements have generally been found impractical for the more complex determinations that require highly sophisticated instrumentation and methodology.

Nevertheless, there has been a constant effort to adapt even the most sophisticated methods for portable use in the field. Such efforts introduce the inevitable challenges of instrument ruggedness, availability of appropriate power supplies, and control of operating parameters.

Field measurement is necessary when samples need to be analysed at their site of occurrence, as, for example, when dealing with unstable analytes; when immediate results are required, such as in detecting chemical or microbiological warfare agents on the battlefield; and when the analyst cannot be present, such as for analyses performed on the surfaces of other planets.

Field Analytical Chemistry and Technology is the first journal to focus on this area, and fills an important niche. It is international and interdisciplinary, publishing refereed, original research papers, short technical notes, reviews and so-called 'Outfield Reports'. The latter articles arise from what is described as "major field analytical projects involving multidisciplinary teams, often operating under highly demanding conditions". Stimulating editorials add to the journal's appeal.

Although the editorial board is international, with members representing nine different countries, and the journal policy is international in scope, 75 per cent of the articles across three issues came from US laboratories.

The journal has a professional, high-quality appearance and is nicely laid out. The period from submission to acceptance is generally less than two months; the lag for actual publication also appears to be short, but is difficult to discern exactly as copies of the journal are identified by volume and issue number within a given year rather than by month of publication. □

Patrick MacCarthy is in the Department of Chemistry and Geochemistry, Colorado School of Mines, Golden, CO 80401, USA.



The Prince of Wales microscope, c. 1755.

A site for sore eyes

Microscopy and Microanalysis

Editor-in-chief Dale E. Johnson
Springer. 6/yr. \$539 (institutional)

L. M. Brown

Microscopes are to modern science what telescopes are to astronomy: a great enabling technology, which has brought about a considerable part of what is now undergraduate physics. Yet microscopy remains a discipline for professionals, and stands quite aloof from mainstream academic science, whether biological or physical.

Technical developments, often carrying in their train consequences of the utmost importance, such as new routes to the fabrication of devices on the nanometre scale, are announced either at specialized conferences or in the pages of *Nature*, but not usually in journals of physics or biology. So journals that cover conference programmes and details of trade exhibitions play a special role. They are sought after by every group that develops or applies microscopical techniques, especially when they concern structure and function of materials, whether biological or inorganic. The latest such publication is *Microscopy and Microanalysis*, the official journal of several major microscopical societies of North and South America. It is affiliated to the major European societies, and makes a fair bid to become the international house journal of microscopists.

It boasts an editorial board filled with eminent American pioneers of microscopy, nicely balanced between those concerned with biological problems and those concerned with the science of materials. North Americans predominate, but there is a smattering of Europeans and South Americans.

The articles are handsomely printed on glossy paper, so that micrographs are excellently reproduced. Their content is typically a mixture of technique and application. About a quarter of each issue is taken up with book reviews, news and commentary, with advance information on a variety of conferences. All aspects of microscopy and microanalysis are covered, although electron microscopy dominates at the moment. There is a lively correspondence section.

Any institution outside America with research programmes in microscopy should have access to a copy of this journal. Active members of the North American community will want to become regular readers. □

L. M. Brown is at the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

Digging up the recent past

International Journal of Historical Archaeology

Editor Charles E. Orser Jr.
Plenum Press. 4/yr. USA \$100, elsewhere \$115 (institutional); \$35 (personal)

Frans Verhaeghe

The archaeologies of the medieval, industrial and particularly early modern periods are relative newcomers, the latter being the most recent addition. Since the 1960s, it has been steadily growing, notably in the Anglo-Saxon world. Elsewhere in Europe, the importance and potential of sixteenth- to eighteenth-century remains are now also slowly being recognized. Journals such as *Post-Medieval Archaeology* and *Historical*

Archaeology (published since 1967 by the UK Society for Post-Medieval Archaeology and the US Society for Historical Archaeology respectively) have created models, but there is room for more international approaches as well as more theoretical and applied research. Hence this new journal.

Judging from the editor's introduction, the journal's target area is the archaeology of any 'historical' period in any region of the planet. It offers medium-sized to more extensive papers, from about 10 to 45 pages long. Some are reprints of less readily accessible papers, but most of the 20-odd papers now published are new, and all bar one focus on the past half-millennium and on the formerly colonized world outside Europe. The contributions are well presented and seem to be published fairly rapidly, in a handy 23 cm by 15 cm standard format, on paper allowing for good drawings and reasonable (though not high-quality) photographs.

A 'Views and Commentaries' section presents innovative and thought-provoking statements on issues confronting archaeologists of historical periods. Reviews and book reviews are not (yet?) offered but comments on published papers occasionally are.

The contents and scope of the first two volumes are very satisfying to anyone interested in ethnic and social issues in archaeology, including the impact of colonization and the role played by different components of material culture in shaping early modern history in many parts of the world.

Apart from the very useful information, I find the most stimulating feature to be the strong direct or indirect emphasis on methodology and theory in archaeology and material culture studies. This contributes not only to a better understanding of our recent past but also to the development of archaeology as a whole. This is why the series gives real value at a very reasonable price. □

Frans Verhaeghe is in the Department of Art History and Archaeology at the Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium.

Free advice for experimenters

Technical Tips Online

Managing editor Mark Patterson
Elsevier. On-line only, free.
<http://tto.trends.com>

Laurie Goodman

Web-surfing scientists have a new reason to catch electronic waves. The fully — and solely — on-line journal *Technical Tips Online* provides quick, easy access to optimal methods for doing what researchers do best: experiments. Two great things about the site: it's free, and it's peer-reviewed. It therefore satisfies every web-surfer's belief that *everything*

on the Internet should be free, and meets the need for high scientific standards for on-line-only resources.

Technical Tips Online is a fusion of a peer-reviewed techniques journal and a methods database. It contains three types of article: peer-reviewed 'Technical Tips' (novel methods or major advances on known methods); non-peer-reviewed 'Protocols' (editorially invited protocols from laboratories with expertise in the method); and non-peer-reviewed 'Application Notes' (company-submitted paid-for reports detailing product use). Editorially selected comments from readers, which are quite fun to read, some with author responses, are also available.

The site is updated fortnightly and is extremely user-friendly. One can browse by category — Polymerase Chain Reaction, Electrophoresis, Microbiology, Purification Methods, Gene Expression, Microscopy, Tissue Culture, and Cloning and Sequencing — or examine only the latest additions. Using its comprehensive search engine is definitely superior to scanning several articles or journals for basic protocols, and then trying to glean modifications in later publications. After a while, however, this site could become equally unwieldy, because it might have to be searched just as extensively to obtain all useful protocol modifications. Given the ease of changing on-line text (which is what makes on-line protocols so appealing), this shouldn't be allowed to occur. Periodic processing of related information by the editors into a single update would increase this site's long-term value.

In this regard, it will be interesting to see how the site develops. Sustaining and properly maintaining any dedicated database is an enormous task, requiring time and money (it is supported by advertisements). All in all, however, if the site is well-maintained and the number of core protocols and useful modifications increases at a rapid rate, this site is well worth a bookmark in my book — er, computer. □

Laurie Goodman is at Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA.

Fuel for thought

Combustion Theory and Modelling

Editors-in-chief Bill Dold and Mitch Smooke

Institute of Physics. 4/yr. USA \$257, elsewhere £127 (institutional); USA \$97, elsewhere £48 (personal)

A. C. McIntosh

Interest in combustion research has grown in recent years because of public pressure to find cleaner, more efficient means of energy production, which will all involve combustion of some kind for generations to come.



A dig at Lime Kiln Dock in London, overlooked by the tower of Canary Wharf.

Because of the nonlinear features of the chemical reactions involved, a whole branch of mathematics has arisen to model the complex behaviour of solids, liquids and gases which are undergoing combustion. This, coupled with great advances in numerical approaches, has led to the formation of a new journal specifically devoted to the "application of mathematical theory, modelling, numerical simulation and experimental techniques to the study of combustion". The inclusion of experimental techniques in its scope is laudable and, it is hoped, will encourage greater involvement of experimentalists in the journal.

Combustion Theory and Modelling has been published since March 1997 and has received papers on a wide area of topics, such as flame stability, turbulent combustion, thermal explosion, compressible reacting flows, detonations and filtration combustion. In that a growing number of authors are making use of the journal, it is evident that it is taking a leading place in the status of archived combustion literature. Furthermore, it is readily available at competitive prices for personal subscription and, most importantly, for library subscription. It is the relative cheapness of library availability that should be most attractive in the present financial climate of universities in the UK and world-wide.

The high profile of theory and modelling in this quarterly is particularly welcome. The study of combustion fundamentals is vital to making progress in our understanding of practical combustion systems. Too often the tendency in research has been to regard theory as either too difficult, or naively to regard computation as somehow having replaced the need for modelling altogether. Advance in any scientific research area is connected to the progress of experiment (numerical or laboratory) coupled with testable theory. That way physical principles are uncovered and light the way for further advance. □

A. C. McIntosh is in the Department of Fuel and Energy, University of Leeds, Leeds LS2 9JT, UK.

Golden age journal

Multibody System Dynamics

Editor Werner Schiehlen

Kluwer Academic Publishers. 4/yr. \$241.50 (institutional), \$85 (personal)

John Hogan

Contemporary mechanics traces its origins to Newton's laws of motion. Traditionally divided into fluid and solid mechanics, the theoretical side in the early part of this century saw ideas such as boundary layer theory give analytic tractability to the otherwise impenetrable continuum equations in these areas. Enormous strides in computer power mean that problems such as the large-scale

computation of global weather patterns can now be addressed. More recently, the qualitative approach crucial to understanding nonlinear dynamics (including 'chaos theory') has been embraced by mechanics to the extent that the subject may be on the verge of a golden age, in which fields previously thought impenetrable can be explored.

The international body in the subject is the International Union of Theoretical and Applied Mechanics (IUTAM). Its president, Werner Schiehlen, is the editor of *Multibody System Dynamics*, which is devoted to one of the most important new fields. Examples of such systems include nearly every man-made object of engineering importance, such as cars, lorries, railway vehicles and robots.

Looking through the articles in this journal, it is clear that the equations of motion can be written down, that computers can produce solutions and that, given enough parameters, sufficient agreement can be found with observation. But something is missing. There seems to be no visceral, intuitive understanding of the behaviour of such systems — with their enormous numbers of degrees of freedom — especially when subject to constraints. Of course this is a highly non-trivial problem, but without a 'feel' for a problem, progress is liable to be restricted. The success and even survival of the journal may well be directly related to this issue.

Nevertheless, Schiehlen is to be congratulated for highlighting this area, and Kluwer applauded for its faith in him. Readers of *Nature* interested in this field should start with Schiehlen's comprehensive review in *Multibody System Dynamics* 1, 149–188 (1997), and then attend IUTAM's congress in Chicago in August 2000. □

John Hogan is in the Department of Engineering Mathematics, Queen's Building, University Walk, Bristol BS8 1TR, UK.

Living on the edge

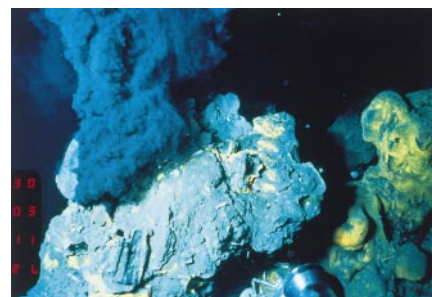
Extremophiles: Life Under Extreme Conditions

Editors Koki Horikoshi, Garo Antranikian, William D. Grant, Terry A. Krulwich and Juergen Wiegand

Springer. 4/yr. \$267, ¥25,000, DM380

John A. Baross

Why do we suddenly need a journal dedicated to micro-organisms that grow under conditions we humans define as extreme? Micro-organisms that grow in saturated concentrations of salt have been studied for almost 80 years. Pressure- and cold-loving bacteria were identified from the deepest ocean trenches during the Danish Galathea expedition in 1952. For at least 20 years we have known of micro-organisms that grow at temperatures close to 100 °C, alkalinities exceeding pH 10, and acidities near pH 0.



The Saracen's Head 'black smoker' vent in the Atlantic Ocean floor.

Yet over the past ten years extremophiles have taken on a life of their own, no longer studied as academic peculiarities. High- and low-temperature enzymes from thermal extremophiles are exploited for their many biotechnological applications, while acidophiles assist in the mining of precious metals from complex minerals. Global phylogenetic trees based on nucleic acid sequences indicate that hyperthermophiles are the oldest of extant organisms, holding potential clues to the origin of life on Earth. Their physiologies suggest that they could grow on other solar bodies or survive transit between them. Such controversial ideas, spurred most recently by evidence that the Jovian moon, Europa, may have a liquid-water ocean, are very much part of the study of extremophiles and the emergent field of 'astrobiology'. Life in extreme environments is in vogue.

Against this backdrop, the creation of a journal dedicated to extremophiles was inevitable. Its appearance is also an acknowledgement of the growing community of researchers dedicated to studying these very exciting organisms from a wide range of disciplinary perspectives. This growth is reflected in the diversity of research papers in the first four issues reporting novel and important new discoveries. The topics range from descriptions of new organisms and physiologies to the characterization of their enzymes and molecular biology. One can also learn that the highest temperature for growth by a pure culture has risen from 110 to 113 °C, that there exist organisms that can oxidize hydrogen gas using ferric-iron as the electron acceptor, and that physiologically diverse heterotrophs exist in the deepest trenches of the ocean.

Overall, the research articles in the first four issues are of high quality. Three issues also offered a very useful review article: I recommend that the editors continue to solicit reviews that keep the community informed of current trends in this rapidly changing field. This attractively produced journal is essential reading for anyone interested in life in extreme environments and its ramifications to other fields of inquiry. □

John Baross is in the School of Oceanography, University of Washington, Box 357940, Seattle, Washington 98195, USA.

Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- γ

Robert T. Nolte^{*†}, G. Bruce Wisely^{*†}, Stefan Westin[‡], Jeffery E. Cobb[§], Millard H. Lambert^{*}, Ritsi Kurokawa[‡], Michael G. Rosenfeld^{||}, Timothy M. Willson[§], Christopher K. Glass[‡] & Michael V. Milburn^{*}

^{*} Department of Structural Chemistry, and [§] Department of Medicinal Chemistry, Division of Chemistry, Glaxo Wellcome Research and Development, Research Triangle Park, North Carolina 27709, USA

[‡] Division of Cellular and Molecular Medicine, and ^{||} Howard Hughes Medical Institute, Department of Medicine, University of California, San Diego, La Jolla, California 92093-0651, USA

[†] These authors contributed equally to this work

The peroxisome proliferator-activated receptor- γ (PPAR- γ) is a ligand-dependent transcription factor that is important in adipocyte differentiation and glucose homeostasis and which depends on interactions with co-activators, including steroid receptor co-activating factor-1 (SRC-1). Here we present the X-ray crystal structure of the human apo-PPAR- γ ligand-binding domain (LBD), at 2.2 Å resolution; this structure reveals a large binding pocket, which may explain the diversity of ligands for PPAR- γ . We also describe the ternary complex containing the PPAR- γ LBD, the antidiabetic ligand rosiglitazone (BRL49653), and 88 amino acids of human SRC-1 at 2.3 Å resolution. Glutamate and lysine residues that are highly conserved in LBDs of nuclear receptors form a 'charge clamp' that contacts backbone atoms of the LXXLL helices of SRC-1. These results, together with the observation that two consecutive LXXLL motifs of SRC-1 make identical contacts with both subunits of a PPAR- γ homodimer, suggest a general mechanism for the assembly of nuclear receptors with co-activators.

The peroxisome proliferator-activated receptor belongs to the family of nuclear receptors that includes the oestrogen receptors, thyroid hormone receptors, retinoic-acid receptors (RARs) and retinoid-X receptors (RXRs)¹. Although many of these receptors have well-studied physiological ligands, such as steroids, thyroxine and retinoids, more than half of the gene family are 'orphan' receptors that, in many cases, are important in development and gene regulation but have not yet been associated with a physiologically relevant ligand². A *bona fide* cellular ligand for PPAR has not yet been found; however, several essential fatty acids, oxidized lipids and prostaglandin derivatives can bind and activate the receptor^{2,3}.

The three known human PPAR subtypes, α , γ and δ , show distinct tissue distributions^{4,5} and are associated with selective ligands^{3,4,6}. PPAR- γ is expressed at high levels in adipose tissue and macrophages and can strongly induce adipogenesis of preadipocyte cells, indicating that an endogenous PPAR- γ ligand may be an important regulator of adipogenesis^{5,7–12}. At submicromolar levels, a class of compounds known as thiazolidinediones (TZDs) activates PPAR- γ and are used pharmaceutically for the treatment of type 2 diabetes because they sensitize tissues to insulin^{13,14}. PPAR- γ may be the biochemical target of TZDs^{6,13}.

PPAR- γ contains the general structural features of other nuclear-receptor-family members, including a central DNA-binding domain and a carboxy-terminal domain that mediates ligand-binding, dimerization and transactivation functions. PPAR- γ must form a heterodimer with RXR to bind DNA and activate transcription. Ligand-dependent transcription requires a highly conserved motif, termed activating function-2 (AF-2), located at the far C terminus of the LBD¹⁵. The structures of the LBDs of RXR, RAR, the thyroid hormone receptor and the oestrogen receptor are similar^{16–19}. The LBDs are folded into three layers of α -helices that allow the formation of a ligand-binding site, which is almost entirely buried within the core of these α -helices. In the structure of the unliganded RXR, the AF-2 motif adopts an α -helical structure that extends away from the LBD. In contrast, the AF-2 helix in the agonist-bound RAR, thyroid hormone receptor and oestrogen

receptor structures is folded against the LBD and the AF-2 hydrophobic residues form part of the ligand-binding pocket. On the basis of these observations, a 'mouse trap' model of receptor activation has been proposed, in which the AF-2 helix closes on the ligand-binding site in response to ligand and establishes a transcriptionally active form of the receptor¹⁸.

There is evidence that nuclear receptors require the ligand-dependent recruitment of co-activator proteins to effectively stimulate gene transcription, which is dependent on allosteric alterations in the AF-2 helical domain^{20,21}. Indeed, all known ligand-dependent nuclear receptors require the AF-2 domain for effective interaction with co-activators^{22–25}. Among the best studied of the co-activators are the p160 proteins, which were initially identified biochemically as ligand-dependent nuclear-receptor-interacting factors and whose genes were cloned and identified as members of a gene family of steroid receptor co-activating factors, referred to as SRC-1/NcoA-1, TIF2/GRIP-1/NcoA-2 and pCIP/ACTR/AIB1 (refs 26–32). The nuclear-receptor-interaction domain of these factors is highly conserved and contains three repeated motifs of the consensus sequence LXXLL (where X is any amino acid), each of which is sufficient for ligand-dependent interaction with several nuclear receptors^{33,34}.

Here we report X-ray structure analysis of the apo-PPAR- γ LBD, and of the ternary structure consisting of the PPAR- γ LBD, the antidiabetic ligand rosiglitazone³⁵, and a region of the SRC-1 nuclear-receptor-interaction domain containing two LXXLL motifs.

Overall structure and refinement

The results of the crystallographic refinement are summarized in Table 1. In both the apo and the ternary complex crystal structures, amino acids 207–476 were modelled into the electron-density map and crystallographically refined. The nomenclature of the secondary-structure elements (Fig. 1a) is based on the RXR LBD crystal structure¹⁶. Most PPAR- γ amino acids were well defined in the electron-density map, except for the loop between helices H2' and

H3, which is the most thermally mobile loop. In the ternary complex, SRC-1 was fitted to the electron density and crystallographically refined for amino acids 628–640 and 684–703. For both the apo and the ternary complex crystals, the same non-crystallographic dimer configuration of PPAR- γ is observed as in the RXR- α and oestrogen receptor- α crystal structures^{16,19}.

The crystal structure of the PPAR- γ LBD reveals a bundle of 13 α -helices and a small four-stranded β -sheet (Fig. 1a). The PPAR- γ LBD is very similar to the overall fold of other nuclear-receptor structures from helix 3 to the C terminus. However, PPAR- γ is unique in its overall tertiary structure and contains an extra helix, designated H2', between the first β -strand and H3. The tertiary placement of H2 is different in PPAR- γ compared with other nuclear-receptor LBD tertiary structures, and provides easier access for ligands. Helices H4, H5 and H8 are tightly packed between helices H1, H3, H7 and H10 at the top half of the LBD (Fig. 1). This structural arrangement seems to be important in forming the ligand-binding site: helices H3, H7 and H10 are anchored to allow them to jut downwards and form the scaffolding for the large cavity of the ligand-binding site in PPAR- γ .

The crystal structures of the apo and ternary complexes of PPAR- γ contain a homodimer of the protein that involves a total surface area of $\sim 1,600 \text{ \AA}^2$. Helix 10 of one monomer forms a coiled-coil interaction with helix 10 of the other monomer, with amino acids F432, A433, L436 and T440 of each monomer contributing to the interaction. Helices 9 and 10 and a loop before helix 9 contribute several salt bridges to the dimer interface; these bridges are formed between residues E418 and K422, E407 and K434, and K438 and D396. There is no evidence that PPAR- γ functions as a homodimer *in vivo*; however, both oestrogen receptor- α and RXR- α use the same homodimer interface as that seen in PPAR- γ . The similarity of the homodimer interfaces in RXR and PPAR indicates that the PPAR and RXR homodimer interface may constitute the heterodimer PPAR–RXR interface for the LBDs. In fact, PPAR and RXR share only 27% sequence identity within their LBDs, but several amino acids that form the strongest interactions at the homodimer

interfaces are conserved in both PPAR and RXR, namely, F432, A433, L436, E418, E407 and K438 of PPAR- γ .

Ligand binding

The apo-PPAR- γ ligand-binding site is a large T-shaped cavity that extends from the C-terminal helix to the β -sheet lying between helices H3 and H6. This domain is mainly hydrophobic and is buried within the bottom half of the LBD (Fig. 1b). One region of the domain is 20 $\text{\AA}$ in length, lies between H3 and the β -sheet, and is parallel to H3. Another cavity, which extends from the β -sheet to the C-terminal AF-2 helix and is 16 $\text{\AA}$ in length, is orthogonal to the T-shaped cavity, behind H3. The total volume of this domain in the apo-PPAR- γ crystal structure is $\sim 1,300 \text{ \AA}^3$ (Fig. 1b). A proposed ligand-entry site in this cavity is apparent from the crystal structure and lies between H3 and the β -sheet. The surface around the entry site comprises several hydrophilic sidechain amino acids, namely, D243, E290, R288 and E295. PPAR- γ is probably unique among other nuclear-receptor LBDs in having this ligand-accessible site^{17–19}.

Rosiglitazone occupies roughly 40% of the ligand-binding site in the ternary complex. In general the ligand is in a U-shaped conformation, wrapping around H3 with its central benzene ring directly behind H3 and the TZD head group and pyridine ring wrapping around H3 (Fig. 2). The TZD head group makes several specific interactions with amino acids in H3, H4, H10 and the AF-2 helix (Fig. 2). The carbonyl groups of the TZD form hydrogen bonds with two histidine residues, H323 and H449. Y473 in the AF-2 helix forms a secondary hydrogen bond with H323. The partly negatively charged nitrogen of the TZD head group is within hydrogen-bonding distance of the Y473 sidechain hydroxyl group. A buried lysine residue, K367, forms another secondary hydrogen bond with the ligand, at residue H449. All of these primary and secondary hydrogen bonds result in a fixed conformation of the TZD head group and of the participating amino acids.

Next to the head group, the sulphur atom of the TZD ring is positioned in a hydrophobic region of the PPAR- γ ligand-binding

Table 1 PPAR- γ LBD structure determinations

Dataset	Apo native ¹	Selenomethionyl	PMA	K ₂ HgCl ₄
Reflections				
Total (N)	41,803	24,798	27,938	24,493
Unique (N)	12,182	11,377	10,760	10,854
R _{sym} (%)	5.5	4.1	6.5	5.8
Resolution (Å)	20.0–3.0	20.0–3.0	20.0–3.1	20.0–3.0
Derivatives				
Phasing power†		1.02	1.42	1.69
R _{scale} (%)‡		7.2	14.5	16.7
Number of sites		12	5	9
Refinement	Apo native ²	SRC-1, rosiglitazone, PPAR- γ		
Space group	C2	P2 ₁ 2 ₁ 2 ₁		
Cell	a = 91.87 Å, b = 62.0 Å, c = 118.3 Å β = 102.66°	a = 52.21 Å, b = 69.06 Å, c = 177.97 Å		
Reflections				
Total (N)	89,611		88,243	
Unique (N)	27,817		25,742	
R _{sym} (%)	5.2		6.1	
Resolution (Å)	60.0–2.2		20.0–2.3	
(% N)	85% (60.0 Å–2.2 Å) 68% (2.3 Å–2.2 Å)		89.7% (20.0 Å–2.3 Å) 53.5% (2.4 Å–2.3 Å)	
Statistics				
R-factors (%) (1σ)	24.6		20.7	
R _{free} (%) (1σ)	31.8		26.4	
r.m.s.d. bond lengths (Å)	0.006		0.008	
r.m.s.d. bond angles (degrees)	1.3		1.6	

* $R_{\text{sym}} = \sum |I_{\text{avg}} - I| / \sum I$.

† Phasing power = $\sum |F_{\text{H}}| / |\sum |F_{\text{PH}}|_{\text{obs}} - |F_{\text{PH}}|_{\text{calc}}|$.

‡ $R_{\text{scale}} = \sum |F_{\text{p}} - F_{\text{p,calc}}| / \sum F_{\text{p}}$.

\$ $R_{\text{factor}} = \sum |F_{\text{p}} - F_{\text{p,calc}}| / \sum F_{\text{p}}$, where F_{p} and $F_{\text{p,calc}}$ are observed and calculated structure factors, R_{free} is calculated for a randomly chosen 5% of reflections and R_{factor} is calculated for the remaining 95% of reflections.

Selenomethionyl, a selenomethionine-incorporated protein; r.m.s.d., root mean square deviation from ideal geometry; %N, percentage of observed reflections used in the crystallographic refinement relative to the total theoretically possible. Apo native¹ and native² are two apo native data sets that were used for MIR structure determination and crystallographic refinement, respectively.

pocket formed by F363, Q286, F282 and L469. The central benzene ring of the ligand occupies a very narrow pocket between C285 and M364. The bridging oxygen atom between the benzene ring and the pyridine ring provides vital geometry for the pyridine ring, which occupies the pocket between H3 and the β -sheet. Rosiglitazone has one chiral atom in the head group that is in the S configuration, even though a racemic mixture was used for crystallization.

α -Substituted carboxylic acids can act as bio-isosteric replacements for the TZD head group, maintaining high affinity binding and receptor activation³⁶. We predict that these carboxylic acids conserve the key interactions with H323 and H449, and that this molecular recognition may be important for the binding of other acidic ligands. Several naturally occurring carboxylic acids, including essential fatty acids, oxidized lipids and prostaglandin J2 metabolites, can bind and activate PPAR- γ at micromolar concentrations². We propose that all of these hydrophobic carboxylic acids bind to PPAR, and that the key interactions involving the histidine residues required for TZD binding are conserved with the acid groups. The other hydrophobic parts of the ligands would be free to interact in a relatively nonspecific manner with the large ligand-binding pocket, resulting in the observed promiscuity of PPAR. In effect, the extra helix, H2', and the altered tertiary placement of the H2 helix define PPAR as a nuclear receptor with a new hormone specificity; it has evolved to bind to multiple natural ligands with moderate affinity.

Determinants of SRC-1 LXXLL interactions

Biochemical studies indicated that two LXXLL motifs were required for cooperative interaction of SRC-1 with nuclear-receptor heterodimers³⁷, and we chose to co-crystallize the PPAR- γ LBD with a region of SRC-1 containing LXXLL motifs 1 and 2 (SRC-1^{623–710}) in the presence of the high-affinity ligand rosiglitazone.

The initial purification and crystallization of the co-expressed PPAR- γ and SRC-1 protein resulted in a molar ratio of 2:1 which was consistent with the observed solution molecular weight (see Methods). The crystal structure shows that the 88 amino acids of SRC-1 bind to a homodimer of PPAR- γ in the presence of ligand. In the ternary complex, one LXXLL motif binds to one molecule of PPAR- γ and the second LXXLL motif binds to the other molecule of PPAR- γ , with the connecting protein spanning the homodimer (Fig. 3b).

Nuclear-receptor AF-2 domains often contain a highly conserved glutamate residue which is important for transactivation but is not required for ligand binding^{38,39}. For example, mutation of this residue in RAR generates a dominant-negative receptor that is transcriptionally silent and can no longer bind to co-activator proteins³⁹. Here, in the ternary complex, the side chain of this residue, E471, forms hydrogen bonds with the backbone amides of K632 and L633 in helical domain HD1 of the LXXLL motif and with the backbone amides of K688, I689 and L690 in helix HD2. In both HD1 and HD2, there is a lysine residue at the amino terminus of the amphiphatic sequence that is solvent-exposed in the crystal structure and does not interact directly with the nuclear receptor. The C-terminal ends of HD motifs interact with the PPAR- γ LBD by a similar hydrogen-bonding network to that seen at the N-terminal ends. A highly conserved lysine, K301, is solvent-exposed in the apo-PPAR- γ structure and seems not to be involved in the folding of the LBD. In the ternary complex, the lysine side chain forms hydrogen bonds with two backbone carbonyls of L636 and T639 in HD1 and L693 and L694 in HD2. There is also biochemical evidence that this lysine residue is important in transcriptional activity and co-activator binding⁴⁰. We propose that E471 and K301 define a 'charge clamp' that allows the orientation and placement of the HD motif into the coactivator-binding site (Fig. 3).

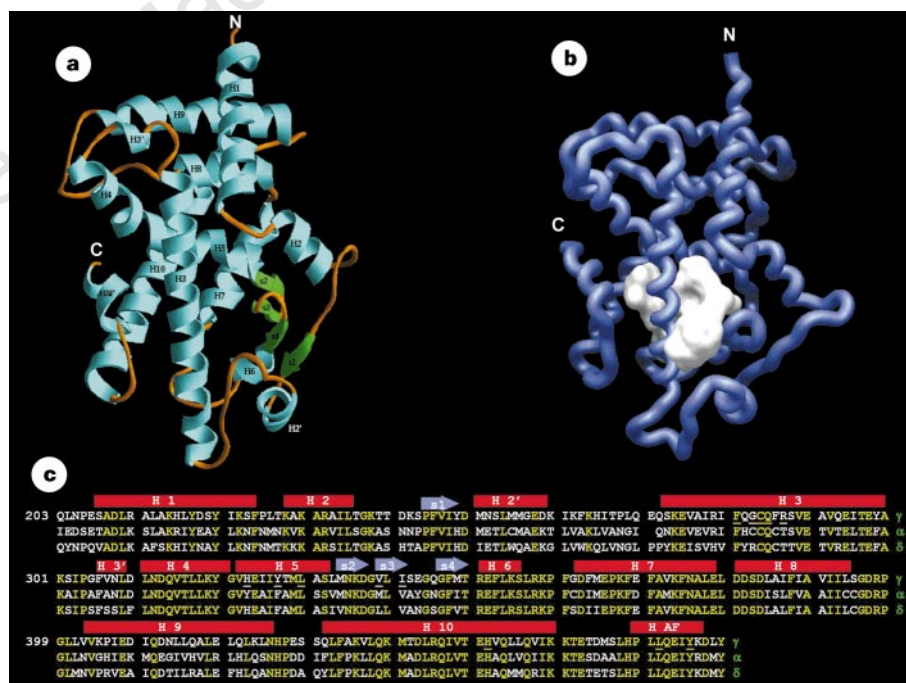


Figure 1 Crystal structure of the apo-PPAR- γ . **a**, Ribbons drawing of the apo-PPAR- γ LBD, amino acids 207–476. The nomenclature of the helices is based on the RXR- α crystal structure¹⁶. α -Helices are light blue; β -strands are green; loops are brown. **b**, A worm backbone tracing of PPAR- γ , with a surface representation of the unoccupied van der Waals space in the ligand-binding site. We determined the 'unoccupied' volume by fitting in virtual atoms that did not occupy the van der Waals surface of the protein. The total unoccupied volume is $\sim 1,300 \text{ \AA}^3$. **c**, Sequence alignment of the human PPAR LBDs. Amino acids that are conserved

between PPARs α , γ and δ are in yellow; the secondary structure that the sequence adopts in PPAR- γ is shown in red boxes for α -helices and blue arrows for β -strands. Residues involved in rosiglitazone binding are underlined. The sequence alignment predicts that several residues involved in direct interactions of PPAR- γ with ligand are not conserved in the other PPAR subtypes, and explains the specificity of TZDs for PPAR- γ : residue H323 is not found in PPAR- α , and Q286 is not found in PPAR- α or PPAR- δ .

The hydrophobic face of the LXXLL helix of SRC-1 packs into a hydrophobic pocket formed between E471 and K301 of PPAR- γ . This pocket is formed from helix-packing interfaces of H3, H4, H5 and the AF-2 helix (Fig. 3). Amino acids L468, L318, T297, Q314, L311, V315, K301 and E471 of PPAR- γ make up most of the binding site for the interacting motifs of SRC-1. The total surface area of the region of interaction between LXXLL and PPAR- γ is $\sim 800 \text{ \AA}^2$ and involves specific salt bridges between the charge clamp on the outside of the binding site and a buried hydrophobic core of the LXXLL motif; L633, L636 and L690, I693 of the two SRC-1 motifs interact hydrophobically with L468 and L318 of the PPAR- γ LBD.

The ternary complex reveals some of the specificity determinants that target the individual HD motifs of SRC-1 to specific nuclear receptors. The charge clamp is presumably not selective in placing the HD motif into this site, although a particular nuclear-receptor dimer may be selective for different co-activators. Selectivity is more likely to reside in the amino acids located N-terminally and C-terminally to the HD motif that are not conserved between co-activators but which do interact with the PPAR- γ LBD. A study of the individual contacts that the HD motifs make with the LBD reveals the role that each residue may play in the recognition process. The residues at positions -3 and -2 of the HD motifs (Table 2) do not appear to make any significant interactions with the LBD; these residues rise up over the charge-clamp residue E471 and away from the surface of the LBD. The amino acid at position -1 fits in a shallow pocket created by P467 and L468 of the AF-2 helix and the +4 residue of the HD motif. There is a preference for large hydrophobic residues at position -1 of the HD motif to allow optimal interaction with PPAR- γ and, in the crystal structure, a lysine residue interacts hydrophobically at this position.

The residues at positions +2 and +3 of the HD motif face out into solution, making no sidechain contacts with the LBD, consistent with the lack of sequence conservation for these positions. The +4 residue lies in a hydrophobic pocket at the surface of the LBD. This leaves the top of the pocket exposed to solvent, perhaps allowing it to easily accommodate residues other than a leucine at this position. The PPAR- γ pocket that encloses the residue at position +5 of the HD motif is built mainly from residues T297, A300, K301, F306, L311 and L318. Large hydrophobic residues, such as methionine or phenylalanine, could fit into this pocket but would require a modest rearrangement of one of the surrounding groups or a shift of the +5 C α position. The residues at the +6 position seem to form a hydrogen bond between the main-chain carbonyl of the +2 residue

and the threonine in HD1 or the glutamine in HD2. Although the +6 residue may be important as a capping residue for helix stabilization, the side chain does not contact the surface of the LBD and is unlikely to convey specificity to the HD. The +7 to +10 residues form no specific sequence, but these residues might interact with other LBD domains and contribute to nuclear-receptor specificity.

Structure and function of AF-2 domains

The two PPAR- γ molecules in the ternary complex have nearly identical C α conformations, with a root mean square (r.m.s.) deviation of 0.76 Å. Although the two molecules in the apo-PPAR- γ homodimer are very similar from the N terminus to amino acid 455, they exhibit entirely different conformations of the AF-2 helix and the preceding loop. In one of the apo-PPAR- γ molecules, the AF-2 helix is folded against the LBD in a conformation that is very similar to that in the ternary ligand-bound complex. In contrast, in the other molecule of the apo-PPAR- γ homodimer the AF-2 helix adopts an extended conformation, projecting slightly away from the LBD. Perhaps, in the apo receptor, the AF-2 helix can assume both an 'active' (AF-2 helix in the liganded state) and an 'inactive' conformation, with the ligand acting to lock the receptor into the active conformation. This scenario would be analogous to the proposed 'mouse trap' model of receptor activation¹⁸. However, a different explanation was suggested by analysis of the interaction of the SRC-1 HD with the ligand-bound receptor. Closer examination of the apo-PPAR- γ structure revealed that the extended AF-2 helix contacted the charge clamp of a crystallographically related PPAR- γ molecule in a different homodimer (Fig. 3f). This crystal packing interaction places the AF-2 helix of one PPAR- γ molecule in an 'inactive' conformation that allows it to bind to the AF-2 helix of another apo-PPAR- γ molecule in a crystallographically related homodimer that is in a conformation similar to that observed for the liganded receptor. E471 and K301 of the 'active' AF-2 helix form hydrogen bonds with the crystallographically related 'inactive' AF-2 helix. The potential generality of this arrangement of AF-2 domains is strengthened by analysis of the apo-RXR- α LBD crystal structure, in which the extended RXR- α AF-2 domain binds the LXXLL-binding pocket of a crystallographically related RXR molecule¹⁶.

Mechanism of ligand activation

Ligand-dependent transcriptional activation by nuclear receptors probably requires the recruitment of co-activator proteins such as

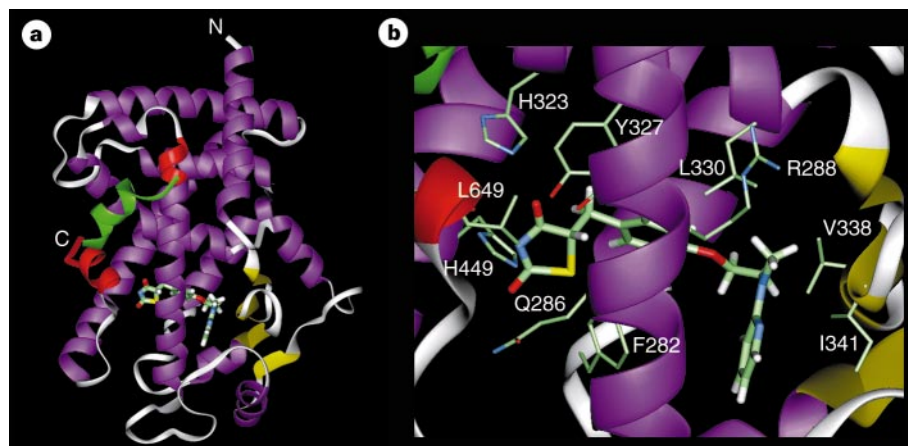


Figure 2 Rosiglitazone binding with the PPAR- γ LBD and SRC-1 in the ternary complex. **a**, Ribbons drawing showing the ternary complex of PPAR- γ LBD, BRL49653, and the LXXLL helix domain of SRC-1. Residues around K301 and E471 that form the 'charged clamp' are red, and the LXXLL SRC-1 helix is green. Rosiglitazone (stick diagram) binds in a deep cavity of the protein and provides a

network of polar interactions that include the AF-2 domain. **b**, The secondary-structure elements are shown as a ribbon drawing, with amino acids involved in ligand binding labelled. Each atom is coloured separately with the same colour scheme: light green, carbon; red, oxygen; blue, nitrogen; yellow, sulphur.

SRC-1 (refs 26–28, 33). Most co-activator proteins contain one or more copies of the LXXLL motif. The ternary structure of PPAR- γ shows that each member of the receptor dimer interacts with a single LXXLL motif of SRC-1. The length and orientation of this helical motif is vital for proper backbone interactions with E471 in the receptor AF-2 domain and with K301 in helix H3 of the LBD. These residues, which are highly conserved among members of the nuclear-receptor family, define a ligand-dependent charge clamp that positions the HD motif of LXXLL to allow the leucine residues to pack tightly into a hydrophobic pocket in the receptor LBD. The residues that define this hydrophobic pocket are also highly conserved among other members of the nuclear-receptor family, suggesting a common mechanism for co-activator binding. The LXXLL motif is also present in several other transcription factors

and in AF-1 domains, indicating that this motif may play a general role in mediating interactions between components of cellular transcriptional machinery.

The observation that two HD motifs of SRC-1 make simultaneous contacts with both LBDs of the PPAR- γ homodimer indicates that the multiple LXXLL motifs within co-activators may mediate cooperative binding to nuclear-receptor dimers. In SRC-1, 52 amino acids separate HD1 and HD2 and 54 amino acids separate HD2 and HD3. There is a similar number of intervening amino acids between the HDs in pCIP and TIF-2. The distance between HD motifs may add more specificity to interactions with nuclear-receptor dimers. Although only two HD motifs would be necessary for cooperative binding of SRC-1 to a nuclear-receptor dimer, perhaps the extra motif (HD3) may allow combinatorial

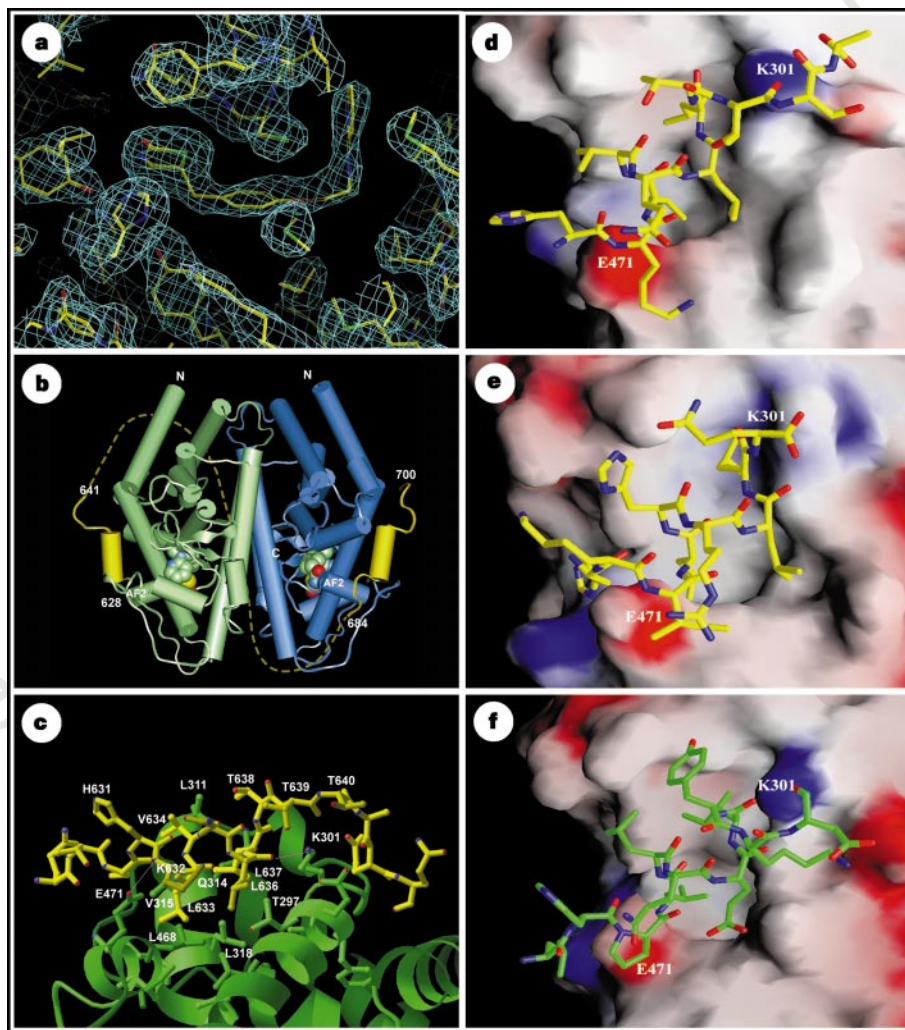


Figure 3 SRC-1 interactions with PPAR- γ . **a**, A sigma-weighted $2F_o - F_c$ omit electron-density map is shown contoured at 1.0σ for the area surrounding the rosiglitazone ligand. **b**, A ribbon drawing of the PPAR- γ LBD dimer and SRC-1, including the ligand rosiglitazone. The two PPAR- γ monomers are blue and green and the two SRC-1 interacting helices are yellow. The structure of SRC-1 was determined from amino acids 628–640 and 684–703 and was crystallographically refined. Very weak electron density from residues 670 to 684 was visible but was not crystallographically refined and is shown as a dashed line. SRC-1 amino acids 642–669 were disordered and not structurally determined. The diagram shows how one SRC-1 molecule, with two interacting domains, forms a complex with a PPAR- γ homodimer. The dashed line connecting the two structurally determined domains of SRC-1 is the proposed connection between these two domains. **c**, The binding of SRC-1 (amino acids 628–642) to the LXXLL-binding site of PPAR- γ . SRC-1 is coloured: yellow, carbon; blue, nitrogen; red, oxygen. The ribbon

backbone of the PPAR- γ LBD is in green. PPAR- γ amino acids binding to the LXXLL helix are also shown in green. **d**, Residues H631–T640 of SRC-1 are coloured as in **c**, with an electrostatic surface of PPAR- γ showing the coactivator-binding site. E471 and K301 side chains result in the red (negative) and blue (positive) charges on the surface of the coactivator-binding site at the N and C termini of the SRC-1 helix, respectively. **e**, Residues H687–E696 of SRC-1 are coloured as in **c**, with an electrostatic surface of PPAR- γ showing the coactivator-binding site. **f**, Amino acids L465–K474 of the PPAR- γ AF-2 helix of one monomer in the apo structure are shown in: green, carbon; blue, nitrogen; red, oxygen, with an electrostatic surface of PPAR- γ showing the coactivator-binding site. E471 and K301 side chains result in the red (negative) and blue (positive) charges on the surface at the N and C terminus of the other PPAR- γ monomer. This figure shows how one monomer in the apo crystal structure orientates its AF-2 helix into the coactivator-binding site of another crystallographically related monomer.

Table 2 AF-2 and HD helix alignments

NR/NCoA	Amino-acid-sequence alignment
Coordinate motif	-432/12345678 LXXLL
SRC-1, HD1	TSHKLVQLTTT
NCoA-2, HD1	GQTKLLQLTTK
P/CIP, HD1	GHKLLQLTCS
SRC-1, HD2	RHKILHRLQEG
NCoA-2, HD2	KHKILHRLQDS
P/CIP, HD2	KHRILHKLQNG
SRC-1, HD3	DHQLRLYLLDKD
NCoA-2, HD3	ENALRLYLLDKD
P/CIP, HD3	NNALRLYLLDRD
PPAR-γ	LHPLLQEIYKDL
Thyroid receptor-α	FPPLFLFVFEHQ
RXR-α	IDTFLMEMLEAP
RAR-γ	MPPLIREMLENP
Oestrogen receptor-α	LYDLLLEMLDAH
Glucocorticoid receptor	FPFMLAEITTNQ
Progesterone receptor	FPFMMSEVIAAQ

Nuclear-receptor (NR) AF-2 helix motifs and co-activator helix motifs are aligned on the basis of the predicted helical structure overlap. The core box of hydrophobic residues that could directly bind to the LXXLL-binding site of the LBD is underlined. There is a high degree of conservation in all of the HD1, HD2 and HD3 boxes, indicating a common function for the serial structure of these domains. The numbering for LXXLL motifs used in the text is shown at the top.

interactions of the LXXLL motifs that are optimal for different receptor dimers.

The observation that the AF-2 helix of an apo receptor can interact with the LBD of a second receptor raises several questions regarding allosteric interactions between nuclear receptors and the mechanisms of ligand-dependent activation. Although the crystal structures of the apo-PPAR-γ and apo-RXR-α LBDs indicate that the AF-2 domains contact the LXXLL-binding pockets of adjacent dimers, molecular modelling indicates that rotation around one amino acid of RXR-α and extension of two amino acids at this position would allow the RXR-α AF-2 domain to bind to the LXXLL-binding pocket of a dimeric partner. Specific partners of RXR in heterodimers, including RAR and the thyroid hormone receptor, allosterically inhibit the binding of ligands to RXR^{41,42}. Allosteric inhibition could be due, in part, to the interaction of the RXR AF-2 helix with the LXXLL-binding pocket of its dimeric partner. There is evidence for this allosteric mechanism and for competition of AF-2 binding with SRC-1 (ref. 37).

Together with the structures of the apo-PPAR-γ LBD and of the ternary complex, these results indicate that ligand-dependent activation may involve displacement of the AF-2 helix of a dimeric partner from the LXXLL-binding pocket. Because the AF-2 domains of many nuclear receptors are strikingly similar, in both length and sequence, to LXXLL motifs, this type of interaction may be relevant to the mechanisms of ligand-dependent activation for many of these proteins (Table 2).

There are several subtle differences between the configurations of the 'active' AF-2 domain in the apo-PPAR-γ LBD structure and the AF-2 domains of the ternary complex that may explain why the binding of ligand would result in a relative preference for binding the LXXLL co-activator motif. Thiazolidinedione binding generates two key hydrogen bonds between H323 and H449 that essentially lock the AF-2 domain into the charged-clamp configuration and alter the sidechain positions in the AF-2 domain that forms the LXXLL-binding pocket. Because the differences between the conformations of the apo 'active' AF-2 domain and the ternary complex AF-2 domain are quite small, it is likely that such changes could be achieved by structurally distinct ligands, indicating a potential basis for the mechanism of action of nuclear-receptor antagonists and partial agonists.

These results suggest that LXXLL motifs are vital in mediating cooperative assembly of co-activator complexes on homodimeric and heterodimeric nuclear receptors. □

Methods

Subcloning and protein purification. Amino acids 176–477 were expressed by subcloning into the T7 *Escherichia coli* expression vector, pRSETa (Invitrogen). The His-tagged plasmid was transformed into BL21(DE3) cells, which were grown at 25 °C, induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG) at $A_{600} = 1.0$ for 1 h, and then collected. The His-tagged PPAR-γ LBD protein was purified with an affinity column of Ni²⁺-coupled Sepharose (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 25 mM imidazole, pH 7.5). The protein was then subjected to limited proteolysis with Arg-C at 4 °C overnight and then purified with a mono-Q-Sepharose column using a salt gradient. The protein was N-terminal sequenced to ensure purity of >95% of the digested PPAR-γ LBD amino acids 183–477. The selenomethionyl protein was prepared in the same way as the wild-type protein.

The SRC1-PPARγ complex was formed by expression in *E. coli* BL21(DE3) of PPAR-γ amino acids 206–477 (plasmid pRSETa; N-terminal His-tag MKKGGHHHHHHG) and SRC-1 amino acids 623–710 (plasmid pACYC184; N-terminal MKK). The transformed cells were grown at 22 °C, induced with 1 mM IPTG at $A_{600} = 1.0$ for 1 h, and then collected. The SRC1-PPARγ complex was purified with an affinity column of Ni²⁺-coupled Sepharose (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 25 mM imidazole, pH 7.5). Two peaks were isolated from the superdex-75 column; these peaks contained PPAR-γ alone and SRC1-PPARγ at a 2:1 molar ratio as quantified using N-terminal sequencing and amino-acid-composition analysis. The 2:1 molar ratio was also consistent with the observed solution molecular weight from gel-filtration analysis and light-scattering measurements.

Crystallization and data collection. The Arg-C-digested PPAR-γ LBD was concentrated to 15 mg ml⁻¹ (20 mM Tris, pH 8.0, 5 mM dithiothreitol (DTT), 100 mM NaCl, 2 mM EDTA) and crystallized using the vapour phase diffusion method by adding equal volume amounts of concentrated protein and a crystallization buffer of 1.4 M sodium citrate and 100 mM HEPES (pH 7.5). We added rosiglitazone to the SRC1-PPARγ complex at a molar ratio of 5:1, concentrated the complex to 10 mg ml⁻¹ (25 mM Tris, pH 8.0, 5 mM DTT, 500 mM ammonium acetate) and crystallized it using the vapour phase diffusion method by adding equal volumes of concentrated protein and a crystallization buffer of 17% PEG 4000, 200 mM ammonium acetate and 20 mM HEPES (pH 7.5). Complete diffraction data were measured at -170 °C with an R-axis II image plate system and an R-axis IV image plate system for the PPAR-γ crystals and the PPARγ-SRC1-rosiglitazone crystals, respectively. Both detectors are mounted on a Rigaku RU-200 rotating anode generator. Data reduction, merging and scaling were done with the programs DENZO and SCALEPACK⁴³.

Phase determination and structure refinement. For the PPAR-γ multiple isomorphous replacement (MIR) structure determination, single isomorphous replacement (SIR) phases from the phenyl mercury acetate (PMA) heavy-atom data were used to calculate a difference map of the selenomethionyl-PPAR-γ data. We identified 12 selenium sites in total, with 6 sites for each individual monomer of the PPAR-γ homodimer. With this information, the non-crystallographic two-fold axis was uniquely determined and used in combination with the MIR phases for electron-density averaging. MIR phases were calculated using PHASES⁴⁴ and improved by solvent flattening and electron-density averaging in program DM^{45,46}. We built the apo-PPAR-γ model using the program Quanta (Molecular Simulations) and refined the model in X-PLOR⁴⁷.

The PPARγ-SRC1-rosiglitazone ternary complex structure was solved by conventional molecular replacement methods using the unliganded PPAR-γ structure, with X-PLOR. Molecular-replacement phases were calculated in the package CCP4 (ref. 46) using SFALL and SIGMAA, and were improved before model building by solvent flattening and density averaging across the PPAR-γ dimer axis with the program DM. Once the SRC1-binding site was determined, residues within 5 Å of the site were excluded from the averaging masks. The resulting electron density was used to identify and build the structure of the SRC-1 protein, rosiglitazone and residues of PPAR-γ around the LXXLL-binding pocket. The unique density for each HD1 and HD2 motif relative to the PPAR homodimer is due to three crystal packing interactions involving SRC-1. We have obtained other holo complexes of thiazolidinediones with PPAR-γ without SRC-1 at a lower resolution than the ternary complex presented, and

they are not significantly different from the ternary structure. The SRC-1 model was built using the software package O (ref. 48), and refined using X-PLOR⁴⁷.

Received 8 May; accepted 29 July 1998.

- Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
- Kliwer, S. A. *et al.* Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors α and γ . *Proc. Natl Acad. Sci. USA* **94**, 4318–4323 (1997).
- Forman, B. M., Chen, J. & Evans, R. M. Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors α and δ . *Proc. Natl Acad. Sci. USA* **94**, 4312–4317 (1997).
- Kliwer, S. A. *et al.* Differential expression and activation of a family of murine peroxisome proliferator-activated receptors. *Proc. Natl Acad. Sci. USA* **91**, 7355–7359 (1994).
- Chawla, A., Schwartz, E. J., Dimaculangan, D. D. & Lazar, M. A. Peroxisome proliferator-activated receptor γ : adipose-predominant expression and induction early in adipocyte differentiation. *Endocrinol.* **135**, 798–800 (1994).
- Willson, T. M. *et al.* The structure-activity relationship between peroxisome proliferator-activated receptor γ agonism and the antihyperglycemic activity of thiazolidinediones. *J. Med. Chem.* **39**, 665–668 (1996).
- Wu, Z., Bucher, N. L. R. & Farmer, S. R. Induction of peroxisome proliferator-activated receptor γ during the conversion of 3T3 fibroblasts into adipocytes is mediated by C/EBP β , C/EBP δ , and glucocorticoids. *Mol. Cell. Biol.* **16**, 4128–4136 (1996).
- Tontonoz, P., Hu, E., Graves, R., Budavari, A. & Spiegelman, B. mPPAR γ 2: tissue-specific regulator of an adipocyte enhancer. *Genes Dev.* **8**, 1224–1234 (1994).
- Tontonoz, P., Hu, E. & Spiegelman, B. M. Stimulation of adipogenesis in fibroblasts by PPAR γ 2, a lipid activated transcription factor. *Cell* **79**, 1147–1156 (1994).
- Ricote, M., Li, A. C., Willson, T. M., Kelly, C. J. & Glass, C. K. The peroxisome proliferator-activated receptor γ is a negative regulator of macrophage activation. *Nature* **391**, 79–82 (1998).
- Tontonoz, P., Nagy, L., Alvarez, J. G. A., Thomazy, V. A. & Evans, R. M. PPAR γ promotes monocyte/macrophage differentiation and uptake of oxidized LDL. *Cell* **93**, 241–252 (1998).
- Brun, R. P. *et al.* Differential activation of adipogenesis by multiple PPAR isoforms. *Genes Dev.* **10**, 974–984 (1996).
- Lehmann, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).
- Hulin, B., McCarthy, P. A. & Gibbs, E. M. The glitazone family of antidiabetic agents. *Curr. Pharm. Des.* **2**, 85–102 (1996).
- Evans, R. M. The steroid and thyroid hormone receptor superfamily. *Science* **240**, 889–895 (1988).
- Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR- α . *Nature* **375**, 377–382 (1995).
- Wagner, R. L. *et al.* A structural role for hormone in the thyroid hormone receptor. *Nature* **378**, 690–697 (1995).
- Renaud, J.-P. *et al.* Crystal structure of the RAR- γ ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681–689 (1995).
- Brzozowski, A. M. *et al.* Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753–758 (1997).
- Horwitz, K. B. *et al.* Nuclear receptor coactivators and corepressors. *Mol. Endocrinol.* **10**, 1167–1177 (1996).
- Glass, C. K., Rose, D. W. & Rosenfeld, M. G. Nuclear receptor coactivators. *Curr. Opin. Cell Biol.* **9**, 222–232 (1997).
- Barettoni, D., Vivanco Ruiz, M. M. & Stunnenberg, H. G. Characterization of the ligand-dependent transactivation domain of thyroid hormone receptor. *EMBO J.* **13**, 3039–3049 (1994).
- Danielian, P. S., White, R., Lees, J. A. & Parker, M. G. Identification of a conserved region required for hormone-dependent transcriptional activation by steroid hormone receptors. *EMBO J.* **11**, 1025–1033 (1992).
- Durand, B. *et al.* Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. *EMBO J.* **13**, 5370–5382 (1994).
- Tone, Y., Chinn, W., T. N., Adams, M. & Chatterjee, V. K. Functional analysis of a transactivation domain in the thyroid hormone beta receptor. *J. Biol. Chem.* **269**, 31157–31161 (1994).
- Voegel, J. J., Heine, M. J. S., Zechel, C., Chambon, P. & Gronemeyer, H. TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. *EMBO J.* **15**, 3667–3675 (1996).
- Torchia, J. *et al.* The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387**, 677–684 (1997).
- Zhu, Y., Qi, C., Calandra, C., Rao, M. S. & Reddy, J. K. Cloning and identification of mouse steroid receptor coactivator-1 (mSRC-1), as a coactivator of peroxisome proliferator-activated receptor γ . *Gene Expr.* **6**, 185–195 (1996).
- Hong, H., Kohli, K., Garabedian, M. J. & Stallcup, M. R. GRIP1, a transcriptional coactivator for the AF-2 transactivation domain of steroid, thyroid, retinoid, and vitamin D receptors. *Mol. Cell. Biol.* **17**, 2735–2744 (1997).
- Anzick, S. L. *et al.* AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* **277**, 965–968 (1997).
- Chen, H. *et al.* Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90**, 569–580 (1997).
- Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414 (1996).
- Voegel, J. J. *et al.* The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *EMBO J.* **17**, 507–519 (1998).
- Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733–736 (1997).
- Keen, H. Insulin resistance and the prevention of diabetes mellitus. *New. Eng. J. Med.* **331**, 1226–1227 (1994).
- Buckle, D. R. *et al.* Non-thiazolidinedione antihyperglycemic agents. 2: α -Carbon substituted β -phenylpropanoic acids. *Bioorg. Med. Chem. Lett.* **6**, 2127–2130 (1996).
- Westin, S. *et al.* Interactions controlling complexes of nuclear-receptor heterodimers and co-activators. *Nature* **395**, 199–202 (1998).
- Feng, X. *et al.* Suprabasal expression of a dominant-negative RXR α mutant in transgenic mouse epidermis impairs regulation of gene transcription and basal keratinocyte proliferation by RAR-selective retinoids. *Genes Dev.* **11**, 59–71 (1997).
- Durand, B. *et al.* Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. *EMBO J.* **13**, 5370–5382 (1994).
- Henttu, P. M. A., Kalkhoven, E. & Parker, M. G. AF-2 activity and recruitment of steroid receptor coactivator 1 to the estrogen receptor depend on a lysine residue conserved in nuclear receptors. *Mol. Cell. Biol.* **17**, 1832–1839 (1997).
- Kurokawa, R. *et al.* Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* **371**, 528–531 (1994).
- Forman, B. M., Umesono, K., Chen, J. & Evans, R. M. Unique response pathways are established by allosteric interactions among nuclear hormone receptors. *Cell* **81**, 541–550 (1995).
- Otwinowski, Z., Isaacs, N. & Burley, S. in *Proc. CCP4 Study Weekend* (ed. Sawyer, L.) 56–62 (SERC Caresbury Lab., Daresbury, 1993).
- Furey, W. S. & Swaminathan, S. in *Methods in Enzymology: Macromolecular Crystallography Part B Vol. 277* (eds Carter, C. & Sweet, R.) 590–620 (Academic, Orlando, 1997).
- Cowtan, K. “DM”: an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newsletter Protein Crystallogr.* **31**, 34–38 (1994).
- Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–776 (1994).
- Brunker, A. X-PLOR Version 3.0: A System for Crystallography and NMR (Yale Univ. Press, New Haven, 1992).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119 (1991).

Acknowledgements. We thank W. Burkhart, M. Moyer and K. Blackburn for protein sequencing and mass spectrometry; S. Blanchard, P. Charifon, T. Consler, S. Jordan, S. Kliwer, J. Lehmann, C. Mohr and E. Xu for discussions; A. Miller for artwork; and K. Milburn for computational assistance.

Correspondence and requests for materials should be addressed to M.V.M. (e-mail: mvm25452@glaxowellcome.com). Coordinates for the apo-PPAR- γ (1PRG) and the PPAR γ -rosiglitazone-SRC1 ternary complex (2PRG) have been deposited with the Brookhaven Protein Database.

A mysterious dust clump in a disk around an evolved binary star system

M. Jura & J. Turner

Department of Physics and Astronomy, University of California, Los Angeles, California 90095-1562, USA

The discovery of planets in orbit around the pulsar PSR1257+12 (ref. 1) shows that planets may form around post-main-sequence stars². Other evolved stars, such as HD44179 (an evolved star which is part of the binary system that has expelled the gas and dust that make the Red Rectangle nebula), possess gravitationally bound orbiting dust disks^{3,4}. It is possible that planets might form from gravitational collapse in such disks⁵. Here we report high-angular-resolution observations at millimetre and submillimetre wavelengths of the dust disk associated with the Red Rectangle. We find a dust clump with an estimated mass near that of Jupiter in the outer region of the disk. The clump is larger than our Solar System, and far beyond where planet formation would normally be expected, so its nature is at present unclear.

The Red Rectangle⁶ is a bright infrared source that exhibits a bipolar optical nebulosity that is illuminated by the A-type star HD44179, an evolved star with total luminosity of $\sim 10^3 L_{\odot}$ (refs 6, 7; here $L_{\odot}$ is the solar luminosity). HD44179 is a 318-day single-lined spectroscopic binary⁸ in which the combined mass of the two stars, M_{bi} , is $\sim 1.5 M_{\odot}$ (here $M_{\odot}$ is the solar mass). It is likely that HD44179 is evolving into a white dwarf, and that the system as a whole is about to become a planetary nebula. Persuasive evidence that much of the circumstellar matter in the Red Rectangle resides in a long-lived orbiting disk is given by the following: the narrow spike (full-width at half-maximum, FWHM, $\sim 2 \text{ km s}^{-1}$) in the circumstellar CO emission line profile³, the peculiar abundances in HD44179^{9,10}, high-spatial-resolution observations at $1.6 \mu\text{m}$ (refs 11, 12) which show a narrow dust lane, the low-surface-brightness extended radio emission at wavelengths of 1.3 cm and 2 cm suggesting



Figure 1 A $15'' \times 13''$ map of the 230-GHz continuum superimposed on a Hubble Space Telescope optical image of the Red Rectangle. The optical image shows the bright central star HD44179 ($m_v = 8.9$ mag, Hipparcos-measured distance of 380 pc). North is to the top, and east is to the left. The beam diameter was $2.8'' \times 2.4''$ at position angle 60° . The primary source lies at right ascension (RA) 06 h 19 min 58.21 s and declination (dec.) $-10^\circ 38' 14.3''$ (equinox 2000.0). East of the primary sources lies the 'secondary source' at RA 06 h 19 min 58.50 s and dec. $-10^\circ 38' 15.0''$ (equinox 2000.0). The contour levels are multiples of 3.5σ each where the noise is 4 mJy per beam.

that grains at least as large as 0.02 cm in radius are present¹³, infrared spectroscopy showing oxygen-rich circumstellar matter¹⁴ around this carbon-rich star¹⁵, and the eccentricity of the orbit of the spectroscopic companion in view of the closeness of the pair of stars^{8,16}.

We used the six-element Owens Valley Millimeter Array during February and March 1996 to perform line and continuum observations at 115 and 230 GHz of the Red Rectangle to map the dust disk. Here we report only the 230-GHz ($1,300\text{-}\mu\text{m}$) continuum results; this is because these results have the higher signal to noise, angular resolution and coverage in the uv plane. The map is shown in Fig. 1. The emission consists of two compact sources within a low surface brightness ($\leq 5 \text{ mJy}$ per beam) extended disk. The millimetre position of the primary source (uncertainty $\pm 0.03''$) is separated by about $0.3''$ from the VLA position measured at 3.6 cm (uncertainty $\pm 0.01''$)¹³. The slight offset between the two positions may be a consequence of the complex geometry and optical thickness at millimetre wavelengths of the inner portions of the dust disk near HD44179¹¹⁻¹³. The secondary source (signal to noise ratio, ~ 10) located at $4.3''$ east of the primary (at a projected distance of $2.5 \times 10^{16} \text{ cm}$ or 1,600 AU) was unexpected because it does not have a counterpart in any existing map of the Red Rectangle at optical⁶ or near-infrared wavelengths^{11,12}, or even at $20 \mu\text{m}$ (ref. 17). Because these previous studies would have revealed an asymmetric disk, the asymmetry in our millimetre and submillimetre observations indicates a distinct, cold source.

We measured fluxes in the primary and secondary sources of 300 mJy and 60 mJy, respectively. The absolute calibration of these fluxes is uncertain to 25%, but the relative fluxes in the two unresolved components are determined to $\sim 10\%$. The total measured flux of $\sim 400 \text{ mJy}$ is consistent with the range of previous measurements¹⁸⁻²⁰. Extrapolating from the extragalactic source counts²¹ at $850 \mu\text{m}$, quite conservatively, there are no more than 10 background sources per degree² at the flux level of the secondary source, and the probability of a background source lying within $4.3''$ of the primary is 5×10^{-5} . Additionally, the secondary lies in the plane of the disk of HD44179; we think it highly likely that the secondary is physically associated with the star.

During three different nights of October and December 1997, we obtained $450\text{-}\mu\text{m}$ observations of the Red Rectangle with the SCUBA camera on the 15-m James Clerk Maxwell Telescope at Mauna Kea, Hawaii in the service observing mode conducted by

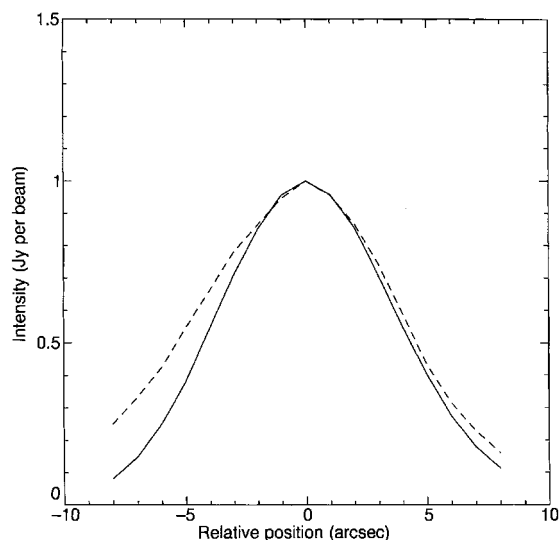


Figure 2 Plot of the normalized intensities at $450 \mu\text{m}$ of east-west scans through the Red Rectangle (solid line) and Uranus used as a calibration source (dashed line). There appears to be a substantial excess to the east in the data for the Red Rectangle. For both the Red Rectangle and Uranus, north-south scans are symmetric at every offset.

G. Sandell. The beam size varied slightly over the different nights; the average FWHM was $8.2''$, too large to resolve fully the secondary source. We show in Fig. 2 plots of east–west scans through the Red Rectangle and the calibrator source, Uranus. At $450\text{ }\mu\text{m}$, the Red Rectangle is markedly stronger in the east. We report a total flux, F_{ν} , from the Red Rectangle of $2.7 \pm 0.3\text{ Jy}$ which is within the range of the previous reported values of $5.7 \pm 0.5\text{ Jy}$ (ref. 19) and $1.46 \pm 0.25\text{ Jy}$ (ref. 20). Depending upon the unknown spatial distribution of the emission to the east, 5–10% of this flux probably arises from the secondary source. With a factor of 2 uncertainty, we estimate $F_{\nu}(450\text{ }\mu\text{m}) \sim 200\text{ mJy}$ for the secondary source.

The result that $F_{\nu}(450\text{ }\mu\text{m})/F_{\nu}(1,300\text{ }\mu\text{m}) = 2\text{--}7$ for the secondary source, lower than the corresponding ratio for the primary source, implies that the temperature of the dust grains, T , is $25 \pm 15\text{ K}$. Because the bulk of the submillimetre emission from the primary source can be fitted¹³ by a black body at a temperature of 50 K and since the secondary source is at least 4 times further from the central illuminating star, the result that $T = 25\text{ K}$ in the secondary source is consistent with being powered by illumination from HD44179, although detailed radiative transfer calculations²² are required for an exact estimate of T .

If D_* denotes the distance from the Sun of the Red Rectangle, if the secondary source is optically thin, if its grains emit at wavelength, λ , with absorption coefficient χ_{ν} , and if the Rayleigh Jeans approximation to the Planck curve is valid, then the mass of dust, M_{du} , is:

$$M_{\text{du}} \approx (F_{\nu} \lambda^2 D_*^2) / (2kT\chi_{\nu}) \quad (1)$$

Assuming that the material is oxygen-rich¹⁴, we adopt $\chi_{\nu}(1.3\text{ mm}) = 1\text{ cm}^2\text{ g}^{-1}$ (ref. 23), but this is uncertain by a factor of 3. With $D_* = 380\text{ pc}$, $T = 25\text{ K}$ and $F_{\nu}(1.3\text{ mm}) = 60\text{ mJy}$, we find with a factor of 4 uncertainty that $M_{\text{du}} \approx 2 \times 10^{30}\text{ g}$, comparable to the mass of Jupiter. If the emission from the secondary source can be approximated by a black body at 25 K , then we extrapolate from the flux at $450\text{ }\mu\text{m}$ to estimate that its total luminosity is $\sim 0.04 L_{\odot}$. In order to emit with this power, a 25 K black body must have a radius of at least $7 \times 10^{14}\text{ cm}$. Because of confusion with the extended low-level disk emission, it is difficult to estimate a maximum size for the secondary, but it appears to have a diameter $\leq 2''$ or a radius $\leq 6 \times 10^{15}\text{ cm}$.

In view of the very peculiar atmospheric abundances in HD44179, at least some of the gas and dust in the circumstellar envelope around the Red Rectangle appear to have separated from each other, and therefore the amount of gas associated with the secondary source is uncertain. A simple model²⁴ for optically thin CO emission³ from a gas with an excitation temperature of 50 K gives a lower bound for the total amount of CO gas in the Red Rectangle of $7 \times 10^{26}\text{ g}$, a factor of 7,000 less than a lower bound to the total grain mass of $5 \times 10^{30}\text{ g}$ derived from the above equation applied to the total 1.3-mm emission. This apparent CO to dust ratio is a factor of $\sim 10^4$ lower than the typical value for a circumstellar envelope²⁵. Although the CO emission may be optically thick as inferred from the observed $^{13}\text{CO}/^{12}\text{CO}$ ratio¹⁹, it is possible that even including H_2 , the secondary source is mostly composed of dust.

We do not have any direct measure of the kinematics of the secondary source; it lies in the plane of the disk of the Red Rectangle at a location where the circular speed is 1 km s^{-1} . Because most of the CO in the entire nebula is moving at less than 2 km s^{-1} with respect to the centre of mass³, either the secondary source is gravitationally bound to the system or it is escaping at not more than a few kilometres per second.

The secondary source presents an unexpected mystery; there is no model to explain the existence of a Jupiter-mass clump of dust lying in the plane of the disk at $\sim 1,600\text{ AU}$ from the central binary. In this cold clump, gravity may be the dominant force. One possible model to explain the clump is that the secondary source is the dust wake produced by a star with a luminosity $\leq 0.04 L_{\odot}$ (implying a mass $\leq 0.5 M_{\odot}$) so that HD44179 is a triple system consisting of the

known spectroscopic binary and a wide-companion third star. If the accumulated dust in the secondary source is gravitationally bound to this putative third star, then it may be contracting into a disk where planets could form. Although the dust extinction could be substantial, deep searches for a stellar counterpart to the secondary source should be undertaken.

Another possibility is that the secondary source is a self-contained clump. In this case, gas thermal or turbulent pressure may provide insignificant support against gravity because of the low gas/dust ratio in the Red Rectangle. Also, because the clump is shielded by the inner disk and because the time when HD44179 will photo-ionize its surroundings is uncertain⁸, the clump may not be dispersed during the planetary nebula phase that the system eventually will enter. Instead, the clump may continue to evolve unperturbed. In this case, most of the internal velocity dispersion, ΔV_{int} , may arise from differential rotation or shear around the central binary; thus $\Delta V_{\text{int}} = 0.5 V_{\text{orb}} \Delta R/R$ where R denotes the distance between the secondary source and HD44179 and V_{orb} denotes the orbital speed for circular motion at R . For the system to be gravitationally bound, $0.5 \Delta V_{\text{int}}^2 M_{\text{du}}$ must be less than $GM_{\text{du}}^2/\Delta R$ which implies, if the secondary source has a circular orbit, that $\Delta R/R \leq 2(M_{\text{du}}/M_{\text{bi}})^{1/3}$. With $R = 2.5 \times 10^{16}\text{ cm}$, $M_{\text{du}} = 2 \times 10^{30}\text{ g}$ and $M_{\text{bi}} = 3 \times 10^{33}\text{ g}$, we find that the clump may be self-gravitating if $\Delta R \leq 4 \times 10^{15}\text{ cm}$, which is within the range of possible values. Currently, although the dust mass of the secondary source is comparable to that of Jupiter, it has a size much larger than that of the Solar System. Its ultimate fate is unknown. $\square$

Received 23 March; accepted 16 June 1998.

- Wolszczan, A. & Frail, D. A. A planetary system around the millisecond pulsar PSR 1257+12. *Nature* **355**, 145–147 (1992).
- Banit, M., Ruderman, M. A., Shaham, J. & Applegate, J. H. Formation of planets around pulsars. *Astrophys. J.* **415**, 779–796 (1993).
- Jura, M., Balm, S. P. & Kahane, C. A long-lived disk around the Red Rectangle? *Astrophys. J.* **453**, 721–726 (1995).
- Waters, L. B. F. M., Waelkens, C., Mayor, M. & Trams, N. R. A model for the 89 Herculis system. *Astron. Astrophys.* **269**, 242–248 (1993).
- Kolykhalov, P. I. & Syunyaev, R. A. The outer parts of the accretion disks around supermassive black holes in galaxy nuclei and quasars. *Sov. Astron. Lett.* **6**, 357–361 (1981).
- Cohen, M. *et al.* The peculiar object HD 44179 “The Red Rectangle”. *Astrophys. J.* **196**, 179–189 (1975).
- Sitko, M. L. Spectral energy distributions of hot stars with circumstellar dust. *Astrophys. J.* **247**, 1024–1038 (1981).
- Waelkens, C., van Winckel, H., Waters, L. B. F. M. & Bakker, E. J. Variability and nature of the Red Rectangle nebula. *Astron. Astrophys.* **314**, L17–L20 (1996).
- Mathis, J. S. & Lamers, H. J. G. L. M. The origin of the extremely metal-poor post-AGB stars. *Astron. Astrophys.* **259**, L39–L42 (1992).
- Waters, L. B. F. M., Trams, N. R. & Waelkens, C. A scenario for the selective depletion of stellar atmospheres. *Astron. Astrophys.* **262**, L37–L40 (1992).
- Roddier, F., Roddier, C., Graves, J. E. & Northcott, M. J. Adaptive optics imaging of proto-planetary nebulae: frosty Leo and the Red Rectangle. *Astrophys. J.* **443**, 249–260 (1995).
- Cruzalebes, P., Tessier, E., Lopez, B., Eckart, A. & Tiphene, D. Diffraction limited near-infrared imaging of the Red Rectangle by bispectral analysis. *Astron. Astrophys. Suppl.* **116**, 597–610 (1996).
- Jura, M., Turner, J. & Balm, S. P. Big grains in the Red Rectangle? *Astrophys. J.* **474**, 741–748 (1997).
- Waters, L. B. F. M. *et al.* An oxygen-rich dust disk surrounding an evolved star in the Red Rectangle. *Nature* **391**, 868–871 (1998).
- Waelkens, C., Van Winckel, H., Trams, N. R. & Waters, L. B. F. M. High-resolution spectroscopy of the central star of the Red-Rectangle nebula. *Astron. Astrophys.* **256**, L15–L18 (1992).
- Artymowicz, P., Clarke, C. J., Lubow, S. H. & Pringle, J. E. The effect of an external disk on the orbital elements of a central binary. *Astrophys. J.* **370**, L35–L38 (1991).
- Hora, J. L., Deutsch, L. K., Hoffmann, W. F. & Fazio, G. G. Mid-infrared imaging of the bipolar nebulae AFGL 618, AFGL 2688, and AFGL 915. *Astron. J.* **112**, 2064–2075 (1996).
- van der Veen, W. E. C. J., Waters, L. B. F. M., Trams, N. R. & Matthews, H. E. A study of dust shells around high latitude supergiants. *Astron. Astrophys.* **285**, 551–564 (1994).
- Greaves, J. S. & Holland, W. S. High mass-loss carbon stars and the evolution of the local $^{12}\text{C}/^{13}\text{C}$ ratio. *Astron. Astrophys.* **327**, 342–348 (1997).
- Walmsley, C. M., Chini, R., Steppe, H., Forveille, T. & Omont, A. 1.3 mm continuum emission from circumstellar envelopes. *Astron. Astrophys.* **248**, 555–562 (1991).
- Smail, I., Ivson, R. J. & Blain, A. W. A deep submillimeter survey of lensing clusters; a new window on galaxy formation and evolution. *Astrophys. J.* **490**, L5–L8 (1997).
- Lopez, B., Tessier, E., Cruzalebes, P., Lefevre, J. & Le Bertre, T. Dust envelope modelling of the Red Rectangle nebula. *Astron. Astrophys.* **322**, 868–877 (1997).
- Pollack, J. B. *et al.* Composition and radiative properties of grains in molecular clouds and accretion disks. *Astrophys. J.* **421**, 615–639 (1994).
- Jura, M., Kahane, C., Fischer, D. & Grady, C. Circumstellar gas in the wide binary HD 188037. *Astrophys. J.* **485**, 341–349 (1997).
- Olofsson, H., Eriksson, K., Gustafsson, B. & Carlstrom, U. A study of circumstellar envelopes around bright carbon stars. *Astrophys. J. Suppl.* **87**, 267–330 (1993).

Acknowledgements. We thank G. Sandell of the JCMT for performing the service observations, the staff at the Owens Valley Radio Observatory, S. Van Dyk and R. Hurt for help with the imaging, and A. Ghez and B. Zuckerman for comments. J.T. was supported by the NSF, and M.J. was supported by NASA.

Correspondence and requests for materials should be addressed to M.J. (e-mail: jura@clotho.astro. ucla.edu).

Unaltered cosmic spherules in a 1.4-Gyr-old sandstone from Finland

Alexander Deutsch*, Ansgar Greshake*†, Lauri J. Pesonen‡ & Pekka Pihlaja‡

* Institut für Planetologie, Westfälische Wilhelms-Universität Münster, Wilhelm-Klemm-Strasse 10, D-48149 Münster, Germany

‡ Geological Survey of Finland, PO Box 96, FIN-02151 Espoo, Finland

Micrometeorites—submillimetre-sized particles derived from asteroids and comets^{1–5}—occur in significant quantities in deep sea sediments^{1,2,4}, and the ice sheets of Greenland^{6,7} and Antarctica^{8,9}. The most abundant micrometeorites are cosmic spherules³, which contain nickel-rich spinels¹⁰ that were crystallized and oxidized during atmospheric entry, therefore recording the oxygen content in the uppermost atmosphere^{10–12}. But the use of micrometeorites for detecting past changes in the flux of incoming extraterrestrial matter, and as probes of the evolution of the atmosphere, has been hampered by the fact that most objects with depositional ages higher than 0.5 Myr show severe chemical alteration². Here we report the discovery of unaltered cosmic spherules in a 1.4-Gyr-old^{13–15} sandstone^{16,17} (red bed) from Finland. From this we infer that red beds, a common lithology in the Earth's history, may contain substantial unbiased populations of fossil micrometeorites. The study of such populations would allow systematic research on variations in the micrometeorite flux from the early Proterozoic era to recent times⁹ (a time span of about 2.5 Gyr), and could help to better constrain the time when the atmospheric oxygen content was raised to its present level^{18–20}.

Micrometeorites represent the whole spectrum of Earth-crossing bodies more completely than do meteorites^{2,5,6,8,9}. Therefore, a systematic study of micrometeorite populations collected in different geological horizons could reveal whether time-related variations in the abundance and type of accreted extraterrestrial matter exist. Such an approach, however, has failed so far, as micrometeorites with high terrestrial residence ages are rarely known. Unique exceptions are cosmic spherules in Eocene² and Jurassic sediments²¹ (about 40 and 190 Myr old, respectively), and the material²² which we describe from the Satakunta sandstone.

The Satakunta sediments (up to 1,800 m thick) fill a graben striking northwest-southeast in southwestern Finland^{14,16,17}. The unit correlates with the Jotnian sandstone¹⁷, overlying the Svecofennian rocks (~1.9–1.8 Gyr old) and rapakivi granites (~1.6 Gyr old) of the Fennoscandian shield. Deposition of the Satakunta sandstone may have already begun during emplacement of rapakivi granites; in contrast the 1,260 ± 10-Myr-old post-Jotnian diabase dykes transecting the sandstone set a stringent age limit for cessation of sedimentation^{13–15}. Sedimentological features of the formation¹⁷ point to rapid deposition in an alluvial environment with only short distances of material transport. Following earlier descriptions^{16,22}, we have sampled the Satakunta sediments at the Murronmäki outcrop (lat. 61° 5' 43.51" N, long. 22° 17' 29.36" E). There, the moderately sorted, medium-grained sandstone is bedded on the decimetre scale, and occasionally interlayered by thin siltstone beds. The purplish red bed containing the cosmic spherules is an arkose with a fine-grained matrix of authigenic quartz, clay minerals and chlorite, in which detrital, slightly rounded quartz and feldspar grains (up to 3 mm sized), as well as a few micas, are embedded. Standard procedures for mineral separation were employed to recover the spherules. From the 60–125 µm heavy mineral fractions

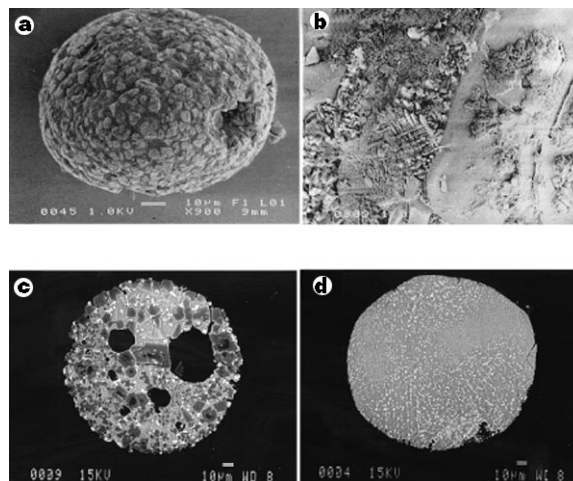


Figure 1 Secondary (a, b), and backscatter (c, d) electron micrographs of cosmic spherules from Murronmäki with (a, c) a porphyritic, and (b, d) a barred-olivine texture. Slight weathering of the spherules in the host sandstone caused preferential removal of the mesostasis yielding a sculptured surface with accentuation of octahedral (a) or skeletal (b) magnetite crystals, which are connected in a typical chain-like fashion (d). The equant olivines in the centre of c are relics with tiny FeNi metal inclusions (bright). a, b, Uncoated samples; field emission SEM JEOL 6300F; operating conditions: 1 kV acceleration voltage, 60 pA beam current. c, d, Polished sections, SEM JEOL 840A; operating conditions: 15 kV acceleration voltage, 3 µA beam current.

of the processed 5 kg sample, 18 strongly magnetic spherules were hand-picked, of which four were sectioned for microanalysis. All the 18 objects are opaque with a metallic shine under reflected light. Most have a spherical morphology (Fig. 1a), yet a few show an elongated shape. Signs of abrasion are totally lacking. We ascribe the excellent morphological preservation to the extraterrestrial spherules having quickly settled in a fluvial environment. Immediate covering by other particles, such as quartz and feldspar, prevented further transportation or re-deposition. Despite the high depositional age of the host lithology, around 1.4 Gyr, the surface of most spherules appears surprisingly unaltered (Fig. 1a, b). Post-depositional processes which resulted in formation of a haematitic pigment in the sandstone matrix did not significantly alter either the extraterrestrial objects or the detrital magnetites which occur in high abundance. Electron microscopy, however, reveals that nearly all the objects suffered light weathering in the sediment, causing sculptured surfaces (Figs 1a, b). The cosmic spherules so far investigated from Murronmäki belong to the S (stony) type³. As shown in Fig. 1, they are “melted micrometeorites” of either (1) the porphyritic (Fig. 1a, c), or (2) the fine-grained barred-olivine type

Table 1 Composition of olivines and matrix in the Satakunta micrometeorites

wt%	1	2	3	4	5	6
SiO ₂	39.8	42.2	42.2	40.5	39.9	41.6
TiO ₂	0.32	n.d.	0.05	n.d.	0.33	0.10
Al ₂ O ₃	5.3	n.d.	n.d.	n.d.	8.9	2.85
Cr ₂ O ₃	0.12	0.48	0.29	0.05	0.06	0.72
FeO	41.5	4.0	2.21	18.6	35.7	21.4
MnO	0.27	0.46	0.41	0.53	0.46	0.25
MgO	6.9	53.0	53.4	40.5	1.83	31.5
NiO	0.16	n.d.	n.d.	n.d.	n.d.	n.d.
CaO	5.2	0.21	0.13	0.03	7.1	0.77
Total	99.57	100.35	98.69	100.21	94.28	99.19
Fayalite (mol%)		4.5	2.9	20.5		

Representative electron microprobe (EMPA) analyses on polished sectioned spherules. Porphyritic textured spherules—no. 1, matrix; olivines, nos. 2, 3, forsteritic cores, and no. 4, rim; barred-olivine spherules—no. 5, matrix (the low total reflects the glassy nature of this material), no. 6, Mg-rich olivine (analysis, taken at the centre of the grain, contains significant contributions of the matrix due to the small grain size of the olivine). n.d., below detection limits of 0.05 wt% for TiO₂, 0.02 wt% for Al₂O₃, and 0.07 wt% for NiO. ‘Matrix’ is defined here as an area which at the resolution of the SEM is free of visible crystal phases. Microprobe used was an JEOL JXA-8600 S, operating at 15 keV and a sample current of 15 nA, at the Institut für Planetologie. Standards: natural silicates and oxides; B and A correction²⁹.

† Present address: Museum für Naturkunde, Humboldt-Universität zu Berlin, Invalidenstrasse 43, D-10115 Berlin, Germany.

(Fig. 1b, d)^{2,3,5}. These textures originate during atmospheric entry of the precursor material at hypervelocities, causing melting and quenching at different cooling rates^{1–8,10,11,23–25}.

The porphyritic spherules are characterized by olivines, up to 20 μm across (Fig. 1c), and small spinels embedded in an apparently glassy, Fe-rich matrix (Table 1). Some olivines contain Mg-rich cores surrounded by an Fe-rich rim (Table 1). In the Mg-rich cores, there is up to 0.2 wt% CaO. In addition, one grain contains several small FeNi metal inclusions (Fig. 1c). These features led us to interpret the olivine cores as relics of the precursor material^{23,24}, which escaped melting during atmospheric entry. The barred-olivine type spherules are also composed of olivines and spinels embedded in an apparently glassy matrix (Fig. 1d), consisting mainly of silica and iron (Table 1). Olivines typically form anhedral to subhedral, 2–5 μm sized crystals with Mg-rich cores (Table 1) rimmed by more Fe-rich compositions; yet relic grains are absent.

The spinels vary widely in grain sizes (maximum, 5 μm) within different regions of one specimen, suggesting a high thermal gradient during cooling. The crystals exhibit skeletal, often cruciform, or octahedral morphologies. In the barred-olivine spherules, the spinels form chain-like, orientated dendritic growths (Fig. 1b, d). Chemically (Table 2), they are magnetites with up to 24 mol% magnesioferrite ($\text{MgFe}_2^{3+}\text{O}_3$). Compositional variation is larger in spinels of the porphyritic spherules, which also contain Al and Cr-rich grains that may represent partial relics of spinels in the precursor material (Table 2).

The textural features, mineral compositions, and bulk chemistry (Table 3) of the Satakunta spherules compare excellently with melted micrometeorites with young terrestrial residence ages^{2–7,23,24}, thus confirming the assessment^{16,22} that they represent micrometeorites. In the porphyritic spherules, some minerals survived the atmospheric entry heating²⁶, indicating that this material experienced lower maximum temperatures compared to the barred-olivine spherules²⁵. The presence of relic MgO-rich olivine cores, high in Ca, and with FeNi metal inclusions, provides the most solid argument for the extraterrestrial origin of the Satakunta spherules. Such olivine compositions are typical of carbonaceous chondrites, relating the precursor material of the porphyritic spherules to this particular class of stony meteorites^{3,23,24}.

Spinel morphology and chemistry further substantiate the extraterrestrial origin. In cosmic spherules, spinels show a restricted range in their $\text{Fe}^{3+}/(\text{total Fe})$ ratio (60–75 at.%), and contain a significant trevorite (NiFe_2O_4) component reflecting the high Ni concentration in the chondritic precursor material^{10,11}. The Satakunta spinels with 1.05–2.65 wt% NiO match spinel compositions in other melted micrometeorites^{7,10,24}; their $\text{Fe}^{3+}/(\text{total Fe})$ ratios fall

exactly in the above given range, with exception of the relic crystals, high in chromium (Table 2).

According to experiments¹¹, the $\text{Fe}^{3+}/(\text{total Fe})$ ratio of spinels which crystallized from melted micrometeorites depends critically on the oxygen fugacity, f_{O_2} , and hence, the altitude of deceleration in the Earth's atmosphere; whereas temperature only plays a subordinate role. In the Satakunta spherules, the $\text{Fe}^{3+}/(\text{total Fe})$ ratios of spinels (Table 2) indicate that f_{O_2} was in the range of 10^{-5} to 7×10^{-4} bar during melting and crystallization^{10,11}. The estimate is corroborated by the composition of the olivine rims ($\sim 70 \text{ mol\%}$ fayalite; Table 1) also serving as experimentally calibrated¹¹ f_{O_2} indicator in melted micrometeorites. In the present atmosphere, such an f_{O_2} corresponds to an altitude of 70–40 km. According to current understanding, a substantial increase in the atmospheric oxygen content occurred during the interval 2.5 to ~ 1.85 Gyr (refs 18–20). In the following late Proterozoic era, the atmospheric oxygen content apparently remained at a value intermediate between modern levels and those of the early Proterozoic^{18,19}, yet higher atmospheric O_2 levels would be consistent with geological data¹⁹. The spinel composition in cosmic spherules of the ~ 1.4 -Gyr-old Satakunta sandstone may favour the latter interpretation, given that the density of the upper atmosphere was similar to the present value.

The depositional age of these sediments exceeds by far that of the hitherto oldest known fossil extraterrestrial material—meteorites in Ordovician sediments²⁷ (~ 460 Myr old). Despite their exceptionally long residence in sediments, the Satakunta micrometeorites still display primary features just like micrometeorites deposited in the Quaternary. The occurrence of cosmic spherules, whose alteration effects are restricted to the outer parts of the rims, in such old sediments opens a new research perspective: a systematic analysis of abundance and type of all spherules in the Satakunta sandstone, and in similar lithologies (red beds) of distinctly different age, could yield a better knowledge on whether the flux of extraterrestrial material has varied with time. For example, ~ 1.4 -Gyr-old red beds occur not only in Finland and Sweden, but are known from the North American craton, the Ukraine, India and Siberia (see, for example, ref. 28). In addition, the presence of unaltered spinels in fossil spherules of different size could provide otherwise unobtainable¹⁹ information on the evolution of the oxygen fugacity in the atmosphere. But if the fossil micrometeorites are found to be restricted to a few layers in the sedimentary rocks, their occurrence may define marker horizons, which are of specific value in fossil-barren red beds. □

Table 2 Composition of spinels in the Satakunta micrometeorites

wt%	1	2	3	4	5	6
TiO ₂	0.25	—	—	0.64	0.14	0.57
Al ₂ O ₃	1.89	2.11	2.36	5.7	1.36	9.0
Cr ₂ O ₃	1.78	1.80	1.74	1.03	0.87	24.5
Fe ₂ O ₃ *	64.4	64.8	64.7	55.5	65.0	27.7
FeO*	23.8	23.3	23.1	35.3	27.7	31.4
MnO	—	0.05	0.11	0.06	0.18	0.21
MgO	6.1	6.3	6.1	0.15	3.7	5.9
NiO	1.78	1.64	1.89	1.62	1.05	0.72
Total	100†	100†	100†	100†	100†	100†
$\text{Fe}^{3+}/(\text{tot. Fe})$ (at.%)	73.0	73.6	73.7	61.1	70.1	46.8

Representative EMPA analyses on polished sectioned spherules; for analytical techniques, see footnotes to Table 1. Analyses 1 to 3—barred olivine; 4 to 6—porphyritic spherules. Grains nos 1 to 3 show a rather uniform composition with $\sim 64 \text{ mol\%}$ magnetite, 20 mol% magnesioferrite, and 4.7 mol% trevorite, and small yet varying amounts of Ti, Cr, Al spinel components. The compositional variation is larger in the porphyritic spherules, ranging from magnetite-magnesioferrite (grain no. 5) to grains rich in either the spinel-hercynite, or the chromite component (grain no. 6); in addition, the ulvöspinel component makes up to 3.2 mol%. The Al and Cr-rich grain no. 6 may represent a partial relic of spinels in the precursor material. Contributions of the host matrix to spinel raw analyses were corrected assuming that all SiO₂ (<10 wt%) comes from the matrix; corrections for other elements were done using elemental ratios as obtained in the matrix.

* Calculated. Molar components of the spinels and atomic $\text{Fe}^{3+}/(\text{total Fe})$ were calculated according to ref. 11).

† Normalized to 100%.

Table 3 Bulk composition of the Satakunta micrometeorites

wt%	Porphyritic spherule A		Barred-olivine spherule B	
	Mean $\pm 1 \sigma$ D	Range	Mean $\pm 1 \sigma$ D	Range
SiO ₂	33.47 $\pm$ 3.41	27.5–35.9	33.32 $\pm$ 0.13	33.3–33.4
TiO ₂	0.18 $\pm$ 0.09	0.06–0.30	0.20 $\pm$ 0.09	0.09–0.27
Al ₂ O ₃	4.74 $\pm$ 0.34	4.23–5.08	3.06 $\pm$ 0.16	2.96–3.35
Cr ₂ O ₃	0.54 $\pm$ 0.25	0.21–0.90	0.49 $\pm$ 0.13	0.35–0.69
FeO ^{tot}	42.54 $\pm$ 4.89	37.8–50.8	39.34 $\pm$ 1.12	38.0–41.0
MnO	n.d.	<0.10	n.d.	<0.05
MgO	15.40 $\pm$ 2.50	13.1–19.5	19.35 $\pm$ 1.54	17.9–21.8
NiO	Irregular	<0.27	0.52 $\pm$ 0.23	0.24–0.86
CaO	2.82 $\pm$ 0.58	1.98–3.52	3.02 $\pm$ 0.19	2.82–3.31
Na ₂ O	Irregular	<0.36	Irregular	<0.72
K ₂ O	Irregular	<0.27	Irregular	<0.03–0.23
SO ₃	0.13 $\pm$ 0.10	<0.24	Irregular	<0.16
Total	99.09–101.2		98.48–101.5	

Energy dispersive analyses (JEOL 840A scanning electron microscope (SEM) with EDS Link AN1000; 20 keV acceleration voltage, 9 nA sample current) on polished sectioned spherules. Standards: natural oxides and silicates; ZAF correction. $N = 5$; scanned area is $20 \times 20 \mu\text{m}$ (A); $80 \times 80 \mu\text{m}$ (B). Relative to the composition of CI chondrites³⁰, the analysed samples are significantly depleted in Ni, Cr, Mn and sulphur, a feature known also from deep sea spherules⁷. The atomic ratios of the spherules normalized to Cl³⁰ cluster at 1 for Ca/Si, range between 0.4 and 0.6 for Fe/Si, and between 0.7 and 1.1 for Mg/Si; the Al/Si ratio, however, exceeds that of CI. As expected, the spherules lack sodium and potassium due to loss by volatilization during atmospheric entry. Minor K₂O and Na₂O contents in some areas may reflect slight weathering in the host sediment.

Received 3 March; accepted 22 June 1998.

- Blanchard, M. B., Brownlee, D. E., Bunch, T. E., Hodge, P. W. & Kyte, F. T. Meteoroid ablation spheres from deep-sea sediments. *Earth Planet. Sci. Lett.* **46**, 178–190 (1980).
- Taylor, S. & Brownlee, D. E. Cosmic spherules in the geologic record. *Meteoritics* **26**, 203–211 (1991).
- Brownlee, D. E., Bates, B. & Schramm, L. The elemental composition of stony cosmic spherules. *Meteoritics Planet. Sci.* **32**, 157–175 (1997).
- Brownlee, D. E. in *The Sea Vol. 7* (ed. Emiliani, C.) 733–762 (Wiley, New York, 1981).
- Brownlee, D. E. Cosmic dust: collection and research. *Annu. Rev. Earth Planet. Sci.* **13**, 147–173 (1985).
- Maurette, M., Jéhanno, C., Robin, E. & Hammer, C. Characteristics and mass distribution of extraterrestrial dust from the Greenland ice cap. *Nature* **328**, 699–702 (1987).
- Robin, E., Christophe Michel-Levy, N., Bourrot-Denise, M. & Jéhanno, C. Crystalline micrometeorites from Greenland blue lakes: their chemical composition, mineralogy and possible origin. *Earth Planet. Sci. Lett.* **97**, 162–176 (1990).
- Maurette, M. et al. A collection of diverse micrometeorites recovered from 100 tonnes of Antarctic blue ice. *Nature* **351**, 44–46 (1991).
- Taylor, S., Lever, J. H. & Harvey, R. P. Accretion rate of cosmic spherules measured at the South Pole. *Nature* **392**, 899–902 (1998).
- Robin, E., Bonté, Ph., Froget, L., Jéhanno, C. & Rocchia, R. Formation of spinels in cosmic objects during atmospheric entry: A clue to the Cretaceous-Tertiary boundary event. *Earth Planet. Sci. Lett.* **108**, 181–190 (1992).
- Gayraud, J., Robin, E., Rocchia, R. & Froget, L. Formation conditions of oxidized Ni-rich spinel and their relevance to the K/T boundary event. *Geol. Soc. Am. Spec. Pap.* **307**, 425–443 (1996).
- Brownlee, D. E., Bates, B. A. & Wheelock, M. M. Extraterrestrial platinum group nuggets in deep-sea sediments. *Nature* **309**, 693–695 (1984).
- Suominen, V. The chronostratigraphy of southwestern Finland, with special reference to Postjotnian and Subjotnian diabases. *Geol. Surv. Fin. Bull.* **304** (1991).
- Amatov, A., Laitakari, I. & Proshin, Ye. in *Explanation to the Map of Precambrian Basement of the Gulf of Finland and Surrounding Area 1:1 mil* (ed. Koistinen, T.) *Geol. Surv. Fin. Spec. Pap.* **21**, 99–113 (1996).
- Neuvonen, K. J. Remanent magnetization of the Jotnian sandstone in Satakunta, SW-Finland. *Bull. Geol. Soc. Fin.* **45**, 23–27 (1973).
- Marttila, E. Satakunnan hiekkakiven sedimentaatio-olosuhteista. (On the sedimentary environment of the Satakunta sandstone.) Thesis, Univ. Turku (1969).
- Kohonen, J., Pihlaja, P., Kujala, H. & Marmo, J. Sedimentation of the Jotnian Satakunta sandstone, western Finland. *Geol. Surv. Fin. Bull.* **369** (1993).
- Des Marais, D. J., Strauss, H., Summons, R. E. & Hayes, J. M. Carbon isotope evidence for the stepwise oxidation of the Proterozoic environment. *Nature* **359**, 605–609 (1992).
- Kasting, J. F. Earth's early atmosphere. *Science* **259**, 920–926 (1993).
- Ohmoto, H. When did the Earth's atmosphere becomeoxic? *Geochim. News* **93**, 12–27 (1997).
- Jéhanno, C., Boclet, D., Bonté, Ph., Castellari, A. & Rocchia, R. Identification of two populations of extraterrestrial particles in a Jurassic hardground of the Southern Alps. *Proc. Lunar Planet. Sci. Conf.* **18**, 623–630 (1988).
- Marttila, E. Mikrometeorititeja Satakunnan kiekakivessä. (Micrometeorites in the Satakunta sandstone.) *Geologi* **20**, 135–136 (1968).
- Steele, I. M. Olivine in Antarctic micrometeorites: Comparison with other extraterrestrial olivine. *Geochim. Cosmochim. Acta* **56**, 2923–2929 (1992).
- Beckerling, W. & Bischoff, A. Occurrence and composition of relict minerals in micrometeorites from Greenland and Antarctica—Implication for their origins. *Planet. Space Sci.* **43**, 435–449 (1995).
- Brownlee, D. E., Bates, B. & Beauchamp, R. H. in *Chondrules and their Origin* (ed. King, E. A.) 122–133 (Lunar and Planetary Inst., Houston, 1983).
- Love, S. G. & Brownlee, D. E. Heating and thermal transformation of micrometeoroids entering the Earth's atmosphere. *Icarus* **89**, 26–43.
- Thorslund, P. & Wickman, F. E. Middle Ordovician chondrite in fossiliferous limestone from Brunflo, central Sweden. *Nature* **289**, 285–286 (1981).
- Goodwin, A. M. *Precambrian Geology; the Dynamic Evolution of the Continental Crust* (Academic, San Diego, 1991).
- Bence, A. E. & Albee, A. L. Empirical correction factors for the electron microanalysis of silicates and oxides. *J. Geol.* **76**, 382–403 (1968).
- Kerridge, J. F. & Matthews, M. S. (eds) *Meteorites and the Early Solar System* Appendix 3 (Univ. Arizona Press, Tucson, 1988).

Acknowledgements. We thank T. Grund, F. Bartschat, U. Heitmann, M. Flucks and D. Kettrup for technical assistance, and A. Bischoff, E. Marttila, A. Putnis, U. Schäfer, and H. Strauss for discussions. Field work was supported by a DAAD-Finnish Academy of Science exchange program (A.D. and L.J.P.) and we acknowledge additional support by DFG.

Correspondence and requests for materials should be addressed to A.D. (e-mail: deusca@uni-muenster.de).

Magnetic trapping of calcium monohydride molecules at millikelvin temperatures

Jonathan D. Weinstein*, Robert deCarvalho*, Thierry Guillet*, Bretislav Friedrich*† & John M. Doyle*

* Department of Physics, † Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, USA

Recent advances^{1–5} in the magnetic trapping and evaporative cooling of atoms to nanokelvin temperatures have opened important areas of research, such as Bose–Einstein condensation and ultracold atomic collisions. Similarly, the ability to trap and cool molecules should facilitate the study of ultracold molecular physics and collisions⁶; improvements in molecular spectroscopy

could be anticipated. Also, ultracold molecules could aid the search for electric dipole moments of elementary particles⁷. But although laser cooling (in the case of alkali metals^{1,8,9}) and cryogenic surface thermalization (in the case of hydrogen^{10,11}) are currently used to cool some atoms sufficiently to permit their loading into magnetic traps, such techniques are not applicable to molecules, because of the latter's complex internal energy-level structure. (Indeed, most atoms have resisted trapping by these techniques.) We have reported a more general loading technique¹² based on elastic collisions with a cold buffer gas, and have used it to trap atomic chromium and europium^{13,14}. Here we apply this technique to magnetically trap a molecular species—calcium monohydride (CaH). We use Zeeman spectroscopy to determine the number of trapped molecules and their temperature, and set upper bounds on the cross-sectional areas of collisional relaxation processes. The technique should be applicable to many paramagnetic molecules and atoms.

Evaporative cooling has proved to be a powerful technique for producing ultracold atoms^{9,15,16}. To employ evaporative cooling, atoms are first loaded into a magnetic trap. Although laser cooling is commonly used to load Li, Na, Rb and Cs, extension of this method to molecules is precluded by their complex internal level structure. (A technique to use multiline Raman cell radiation to optically cool alkali-metal dimers has been proposed, but not yet implemented¹⁷.) Hydrogen has also been trapped but is the only atoms that can be loaded by surface thermalization, as its very low binding energy to liquid helium (1 K) is unique. A molecule would be strongly adsorbed by such a cold surface. A different approach to produce cold molecules for trapping is photoassociation of laser-cooled alkali-metal atoms^{18,19}. This technique has recently been demonstrated²⁰, producing Cs₂ molecules at a translational temperature of 300 μ K. This is sufficiently cold that it should be possible to trap these molecules with far-off-resonance optical trapping

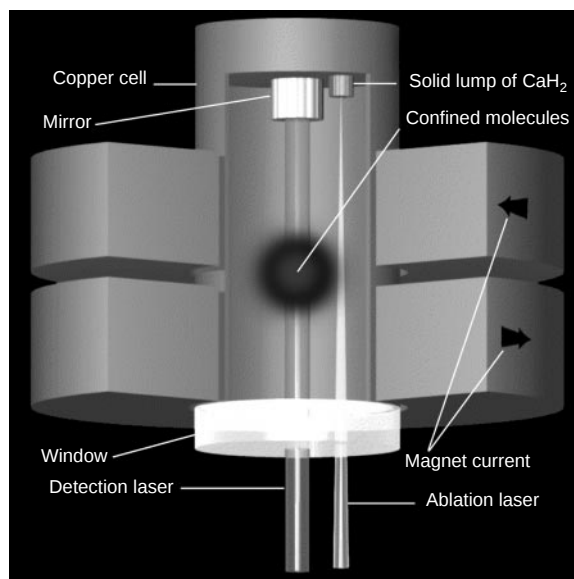


Figure 1 Cutaway diagram of the experimental apparatus. The copper cell is anchored to the mixing chamber of a dilution refrigerator. The magnet is immersed in liquid helium. A vacuum space isolates the relatively warm (4 K) magnet from the cold (300 mK) cell. The superconducting magnet coils are arranged in the anti-Helmholtz configuration (currents travel in opposite directions) and generate a spherical quadrupole magnetic trap up to 3 T deep. For the 1 μ_B magnetic moment of the ground state of CaH, this results in a trap depth of 2 K. Detection and ablation lasers enter through three borosilicate windows (not shown) at 300 K, 77 K and 4 K before passing through the fused silica window at the bottom of the cell. Fluorescence from the trapped molecules (induced by the detection laser) is collected outside the 300 K window by a photomultiplier tube.

techniques^{21–24}. We are aware of work where internally excited, translationally cold Cs₂ molecules have been produced and trapped in this manner²⁵.

In an attempt to broaden the range of atoms that can be trapped, and to trap molecules, we developed a general trap loading technique¹² that relies only on elastic collisions with a cold buffer gas of helium. A cryogenically cooled buffer gas thermalizes atoms or molecules to temperatures less than the depth of the trap.

We chose to trap CaH because it is a paramagnetic molecule that has been previously studied and analysed at zero field. CaH has a magnetic moment of 1 Bohr magneton (1 μ_B) in the ground state. Also, its relatively simple structural and spectroscopic properties aid in detection. Finally, it was straightforward to produce CaH via laser ablation, a process particularly well suited to our cryogenic environment.

The principles of buffer-gas loading and the specifics of our apparatus are given elsewhere^{12,13}. Briefly, our trap is a magnetic potential with a magnetic-field minimum in free space. This field minimum attracts molecules in the low-field-seeking state (that is, those with their magnetic moments orientated anti-parallel to the magnetic field), and repels strong-field-seeking molecules. Our magnetic trap is a spherical quadrupole field (see Fig. 1). Although particles in such a trap are subject to non-adiabatic spin flip (Majorana) losses near the trap centre, our trap parameters are such that these losses are negligible during our observation time^{26,27}.

CaH is produced, thermalized, and trapped within a copper cell inside the bore of the magnet. The cell is maintained at cryogenic temperatures (typically 300 mK) by a dilution refrigerator. The cell is filled with ³He gas (typically at densities of $2 \times 10^{16} \text{ cm}^{-3}$). CaH is created via laser ablation of a solid sample of CaH₂ near the edge of the trap potential. A fused-silica window on the bottom of the cell permits optical access for the ablation beam, as well as a probe laser beam used to detect the molecules. Typically, a 10 mJ laser pulse (5 ns long) at 532 nm is used to ablate solid CaH₂ and produce molecular CaH. The CaH molecules diffuse through the helium gas,

and elastic collisions between the CaH and the helium gas thermalize the translational temperature of the molecules to that of the helium. The cold molecules in the weak-field-seeking states are confined by the magnetic fields and trapped. However, because these trapped molecules are thermally distributed, they evaporate over the edge of the trap at a rate dependent on the ratio of the trap depth to the translational temperature of the molecules.

The CaH molecules are detected by laser fluorescence spectroscopy. They are excited on the $|B^2\Sigma, \nu' = 0\rangle \leftarrow |X^2\Sigma, \nu'' = 0\rangle$ transition at 634 nm (refs 28, 29). The fluorescence emitted on the $|B^2\Sigma, \nu' = 0\rangle \rightarrow |X^2\Sigma, \nu'' = 0\rangle$ transition at 690 nm is detected with a photomultiplier tube (here ν is the vibrational quantum number of the CaH state and ' ' ' designates the excited (ground) state). Colour filters are used to block the scattered probe radiation (634 nm) and pass the frequency-shifted fluorescence (690 nm). The Franck–Condon factor for the $\nu' = 0 \rightarrow \nu'' = 1$ decay has been calculated³⁰ to be 0.028. In order to determine the absolute number of molecules, we calibrate our fluorescence spectroscopy with absorption spectroscopy³¹. We found that under certain conditions 10^{10} ground-state CaH molecules could be produced in a single ablation pulse.

The ground-state molecules are detected on the $|N' = 1, J' = 3/2, \nu' = 0\rangle \leftarrow |N'' = 0, J'' = 1/2, \nu'' = 0\rangle$ transition (N, J , and M_J are the rotational, angular momentum, and projection of angular momentum quantum numbers, respectively). In a magnetic field, this transition is observed to split into multiple features. The detection and analysis presented here was performed on the two slightly broadened features which undergo opposite frequency shifts in a magnetic field. The measured shifts of these features are linear in field and indicate a small difference (0.02 μ_B) in the magnetic moments of the ground and excited states. A theoretical analysis (B.F. *et al.*, manuscript in preparation) indicates that the feature shifted towards higher frequencies is the transition from the $|N'' = 0, J'' = 1/2, M_J'' = 1/2\rangle$ low-field-seeking state to the $|N' = 1, J' = 3/2, M_J' = 3/2\rangle$ excited state, and the feature that shifts towards lower frequencies is the transition from the $|N'' = 0, J'' = 1/2, M_J'' = -1/2\rangle$ high-field-seeking state to the $|N' = 1, J' = 3/2, M_J' = -3/2\rangle$ excited state. The magnetic splitting observed is largely due to a perturbation of the excited state by a nearby $A^2\Pi, \nu = 1$ state, which is 13 cm^{-1} above the $|N' = 1, J' = 3/2\rangle$ state^{28,29}.

Figure 2 displays spectra of these features taken at various times after the ablation pulse. The magnetic trap depth is 3 T, and the cell temperature is $275 \pm 10 \text{ mK}$ before the ablation pulse. Knowledge of the frequency shift as a function of field (in addition to the field

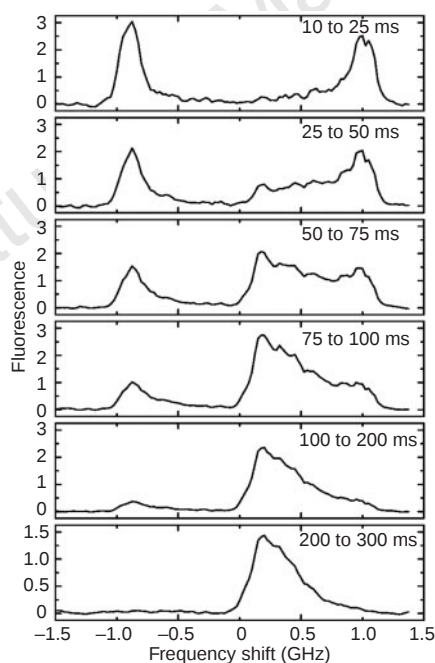


Figure 2 Time evolution of the CaH spectrum. The data were obtained by firing a single ablation pulse and monitoring the fluorescence as a function of time at a single probe frequency. The process was repeated at different frequencies, and the fluorescence averaged over the specified time window in order to obtain the spectra. Times quoted are relative to the ablation pulse. These spectra reveal that high-field-seekers (at negative frequency shifts) quickly leave the trap. The trapped low-field-seekers (positive frequency shifts) are confined and compressed towards the centre of the trap.

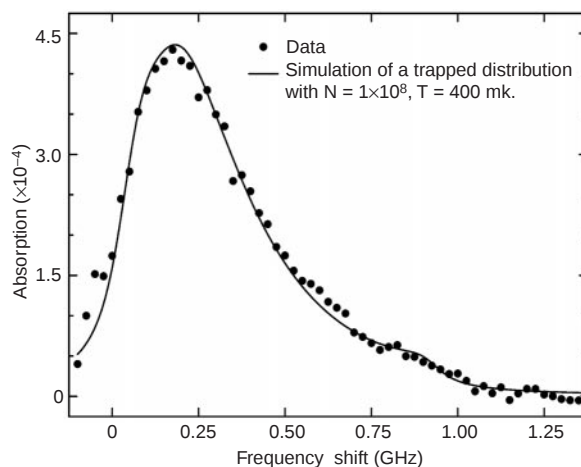


Figure 3 Fit of CaH spectrum to a thermal distribution of trapped atoms. Data were taken with a 3 T trap depth and a cell temperature of 300 mK before ablation, and were taken from 200 ms to 300 ms after the ablation pulse. This fit determined the number and temperature of trapped CaH molecules to be $N = 1 \times 10^8$ and $T = 400 \text{ mK}$.

distribution, selection rules, and probe beam geometry) allows the determination, from the data, of the distribution of the low-field and high-field seeking states of CaH in the magnetic field. Because the shift is linear in field (over the range of fields in our trap), the magnitude of the frequency shift of the observed transition of a CaH molecule is simply proportional to the field it is in. Analysis of our data confirms that the transitions originating from the low-field-seeking state and the high-field-seeking state are shifted to lower frequencies (the left-hand side of our spectra) and higher frequencies (the right-hand side), respectively.

As can be seen in Fig. 2, the spectrum is initially nearly symmetric, indicating that at short times after the production of CaH, the distribution of the high-field-seekers is the same as that of the low-field-seekers. At these short times, there are peaks at high spectral shift, indicating a greater number of molecules in high field than in low field. This is due to a greater volume of high-field regions within the cell. After the ablation pulse, the high-field-seeking CaH molecules (the left-hand spectral feature) rapidly leave the trap. This decay is exponential with the same time constant as that of CaH produced with the magnetic trap off (under otherwise similar conditions). We believe this to be the timescale on which the CaH diffuses to the walls, where it sticks.

In sharp contrast to the high-field-seeking distribution, the low-field-seeking distribution changes shape dramatically, and exists for a longer time. The movement of the right-hand-side spectral feature towards lower shifts indicates that the low-field-seeking CaH molecules are moving in towards lower field. The CaH molecules move from their initial spatial distribution to a thermal, trapped distribution, increasing their density by over an order of magnitude. This occurs on the same diffusion timescale as discussed above, as the CaH moves through the helium to reach its equilibrium distribution.

By fitting the observed spectrum to simulated spectra of thermal distributions of trapped molecules, we determine the number and temperature of trapped CaH. Figure 3 shows a spectrum of trapped CaH along with a fit to a simulated spectrum of a thermal distribution of trapped CaH. From this fit, we determine there are 1×10^8 CaH molecules trapped at a temperature of 400 ± 50 mK. The trapped molecules are at a density of $8 \times 10^7 \text{ cm}^{-3}$ and are lost from the trap with an exponential decay constant of 0.5 s. They can be observed for longer than 2 s. Comparable production conditions without the magnetic trap produced 3×10^8 CaH molecules at densities of only $4 \times 10^6 \text{ cm}^{-3}$; about 30% of the molecules produced are trapped with a density compression factor of 20. (Quoted numbers and densities are accurate to within a factor of two.)

In addition to the trapping and study of CaH molecules in the ground vibrational state ($v=0$), we were also able to detect $v=1$ molecules at zero field. These vibrationally excited molecules (at an energy of 10^3 cm^{-1} higher than the ground state) are out of thermal equilibrium with the translational degrees of freedom. From these data, we were able to determine the vibrational relaxation cross-section between CaH and He as $\sigma_v < 10^{-18} \text{ cm}^2$ at 0.3–1 K. This cross-section is sufficiently low that it should be possible to magnetically trap these molecules via buffer-gas loading. However, production efficiency of the $v=1$ state was significantly less (typically by a factor of 10–100) than for the $v=0$ state. Due to this low signal, we did not attempt to observe trapping of vibrationally excited molecules.

Vital to the success of buffer gas loading is a small cross-section for 'spin-flip' collisions between the buffer gas (helium) and the molecules to be trapped. A low-field-seeking molecule is in a higher energy state than a high-field-seeking molecule. It has the possibility of relaxing to the high-field-seeking state during a collision. In order for buffer gas loading to work efficiently, this spin relaxation cross section (σ_s) must be much smaller than the elastic cross-section (σ_e). The condition $\sigma_e \gg \sigma_s$ allows kinetic thermalization to take place before spin relaxation. From the measured lifetimes of trapped CaH, we have determined the CaH–He spin relaxation cross-

section to be $\sigma_s < 10^{-18} \text{ cm}^2$. From the data on diffusive loss (to the walls) of CaH molecules at zero field, the CaH–He elastic cross-section is $\sigma_e \geq 10^{-14} \text{ cm}^2$, resulting in a ratio $\sigma_e/\sigma_s < 10^4$. As thermalization takes place in ~ 100 collisions, the condition of small σ_s is well satisfied.

Removal of the buffer gas after trapping should lead to evaporative cooling. The cryopumping procedure we used^{13,14} with Cr and Eu takes ~ 10 s, which is longer than we can currently observe trapped CaH. Improvements in the experiment such as a deeper trap, more efficient CaH production, or higher detection sensitivity would lead to longer observation times, allowing time to pump out the helium. Alternatively, more rapid pumping methods could be developed.

Because our technique relies only on elastic collisions with the buffer gas and on the magnetic state of the species, it should be applicable to many paramagnetic atoms and molecules. Implementation of evaporative cooling could lead to the production of large numbers of ultracold molecules. Immediate applications of this work include high-resolution spectroscopy and studies of cold collisions. Future opportunities may include production of quantum degenerate molecular gases and improved searches for permanent electric dipole moments of elementary particles. □

Received 16 July; accepted 17 August 1998.

- Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (1995).
- Davis, K. B. *et al.* Bose-Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
- Bradley, C. C., Sackett, C. A. & Hulet, R. G. Bose-Einstein condensation of lithium: observation of limited condensate number. *Phys. Rev. Lett.* **78**, 985–989 (1997).
- Yu, I. A. *et al.* Evidence for universal quantum reflection of hydrogen from liquid ^4He . *Phys. Rev. Lett.* **71**, 1589–1592 (1993).
- Wineland, D. J., Wieman, C. E. & Smith, S. J. (eds) *Proc. 14th Int. Conf. (Am. Inst. Phys., New York, 1995)*.
- Balakrishnan, N., Forrey, R. C. & Dalgarno, A. Quenching of H_2 vibrations in ultracold ^3He and ^4He collisions. *Phys. Rev. Lett.* **80**, 3224–3227 (1998).
- Hinds, E. A. & Sangster, K. in *Time-Reversal—The Arthur Rich Memorial Symposium* (eds Skalsey, M., Bucksbaum, P. H., Conti, R. S. & Gidley, D. W.) 77–83 (Am. Inst. Phys., New York, 1993).
- Migdal, A. L., Prodan, J. V., Phillips, W. D., Bergeman, T. H. & Metcalf, H. J. First observation of magnetically trapped neutral atoms. *Phys. Rev. Lett.* **54**, 2596–2599 (1985).
- Ketterle, W. & Van Druten, N. J. Evaporative cooling of trapped atoms. *Adv. At. Mol. Opt. Phys.* **37**, 181–236 (1996).
- Hess, H. F. *et al.* Magnetic trapping of spin-polarized atomic hydrogen. *Phys. Rev. Lett.* **59**, 672–675 (1987).
- van Roijen, R., Berkhout, J. J., Jaakkola, S. & Walraven, J. T. M. Experiments with atomic hydrogen in a magnetic trapping field. *Phys. Rev. Lett.* **61**, 931–934 (1988).
- Doyle, J. M., Friedrich, B., Kim, J. & Patterson, D. Buffer-gas loading of atoms and molecules into a magnetic trap. *Phys. Rev. A* **52**, R2515–R2518 (1995).
- Kim, J. *et al.* Buffer-gas loading and magnetic trapping of atomic europium. *Phys. Rev. Lett.* **78**, 3665–3668 (1997).
- Weinstein, J. D. *et al.* Magnetic trapping of atomic chromium. *Phys. Rev. A* **57**, R3173–R3175 (1998).
- Hess, H. F. Evaporative cooling of magnetically trapped and compressed spin-polarized hydrogen. *Phys. Rev. B* **34**, 3476–3479 (1986).
- Masuhara, N. *et al.* Evaporative cooling of spin-polarized atomic hydrogen. *Phys. Rev. Lett.* **61**, 935–938 (1988).
- Bahns, J. T., Stwalley, W. C. & Gould, P. L. Laser cooling of molecules: a sequential scheme for rotation, translation, and vibration. *J. Chem. Phys.* **104**, 9689–9697 (1996).
- Cote, R. & Dalgarno, A. Mechanism for the production of vibrationally excited ultracold molecules of $^7\text{Li}_2$. *Chem. Phys. Lett.* **279**, 50–54 (1997).
- Band, Y. B. & Julienne, P. S. Ultracold-molecule production by laser-cooled atom photoassociation. *Phys. Rev. A* **51**, R4317–R4320 (1995).
- Fioretti, A. *et al.* Formation of cold Cs_2 molecules through photoassociation. *Phys. Rev. Lett.* **80**, 4402–4405 (1998).
- Askar'yan, G. A. Effects of the gradient of a strong electromagnetic beam on electrons and atoms. *Sov. Phys. JETP* **15**, 1088–1090 (1962).
- Ashkin, A. Trapping of atoms by resonance radiation pressure. *Phys. Rev. Lett.* **40**, 729–732 (1978).
- Miller, J. D., Cline, R. a. & Heinzen, D. J. Far-off-resonance optical trapping of atoms. *Phys. Rev. A* **47**, R3567–R3570 (1993).
- Friedrich, B. & Herschbach, D. Alignment and trapping of molecules in intense laser fields. *Phys. Rev. Lett.* **74**, 4623–4626 (1995).
- Takekoshi, T., Patterson, B. M. & Knize, R. J. *Phys. Rev. Lett.* (submitted).
- Petrach, W., Anderson, M. H., Ensher, J. R. & Cornell, E. A. Stable, tightly confining magnetic trap for evaporative cooling of neutral atoms. *Phys. Rev. Lett.* **74**, 3352–3355 (1995).
- Kim, J. *Buffer-gas Loading and Magnetic Trapping of Atomic Europium*. Thesis, Harvard Univ. (1997).
- Berg, L.-E. & Klynnning, L. Rotational analysis of the A-X and B-X band systems of CaH. *Physica Scripta* **10**, 331–336 (1974).
- Martin, H. Laser spectroscopic investigations of the red band systems of CaH. *J. Mol. Spectrosc.* **108**, 66–81 (1984).
- Leininger, T. & Jeung, G. Ab initio calculation of rovibronic transition spectra of CaH. *J. Chem. Phys.* **103**, 3942–3949 (1995).
- Berg, L.-E., Ekvall, K. & Kelly, S. Radiative lifetime measurement of vibronic levels of the $B^2\Sigma^+$ state of CaH by laser excitation spectroscopy. *Chem. Phys. Lett.* **257**, 351–355 (1996).

Acknowledgements. We thank R. Field for discussions on the theory of the Zeeman structure of CaH. This Letter is based on work supported by the US NSF; J.D.W. is supported by an NSF graduate research fellowship.

Correspondence and requests for materials should be addressed to J.M.D.

Highly efficient phosphorescent emission from organic electroluminescent devices

M. A. Baldo*, D. F. O'Brien*, Y. You†, A. Shoustikov†, S. Sibley†‡, M. E. Thompson† & S. R. Forrest*

* Center for Photonics and Optoelectronic Materials, Department of Electrical Engineering and the Princeton Materials Institute, Princeton University, Princeton, New Jersey 08544, USA

† Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA

The efficiency of electroluminescent organic light-emitting devices^{1,2} can be improved by the introduction³ of a fluorescent dye. Energy transfer from the host to the dye occurs via excitons, but only the singlet spin states induce fluorescent emission; these represent a small fraction (about 25%) of the total excited-state population (the remainder are triplet states). Phosphorescent dyes, however, offer a means of achieving improved light-emission efficiencies, as emission may result from both singlet and triplet states. Here we report high-efficiency ($\geq 90\%$) energy transfer from both singlet and triplet states, in a host material doped with the phosphorescent dye 2,3,7,8,12,13,17,18-octaethyl-21H,23H-porphine platinum(II) (PtOEP). Our doped electroluminescent devices generate saturated red emission with peak external and internal quantum efficiencies of 4% and 23%, respectively. The luminescent efficiencies attainable with phosphorescent dyes may lead to new applications for organic materials. Moreover, our work establishes the utility of PtOEP as a probe of triplet behaviour and energy transfer in organic solid-state systems.

When the absorption spectrum of the acceptor (dye) overlaps the emission spectrum of the donor (host), efficient energy transfer from the host to the dye can occur via a singlet-allowed, induced-dipole coupling between the molecular species. Hence, for a fluorescent emitter, the maximum external quantum efficiency (photons extracted in the forward direction per electron injected) is^{4,5}:

$$\Phi_{\text{el}} = \chi \Phi_{\text{fl}} \eta_{\text{r}} \eta_{\text{e}}$$

The fraction of charge carrier recombinations in the host resulting in singlet excitons is χ , which from spin statistics is presumed to be $\sim 1/4$, Φ_{fl} is the photoluminescent efficiency of the dye, η_{e} is the fraction of emitted photons that are coupled out of the device, and η_{r} is the fraction of injected charge carriers that form excitons. As both the recombination and fluorescent efficiencies can approach unity for an optimized device, the efficiency is primarily limited by coupling losses and a restriction to singlet excitons imposed by spin conservation in the induced-dipole energy-transfer process.

Although the output coupling of photons can be increased by using shaped substrates⁶, further efficiency improvements require that both singlet and triplet excited states contribute to luminescence. It has been proposed that intersystem crossing in lanthanide complexes may achieve this with an intramolecular energy transfer from a triplet state of the organic ligand to the $4f$ energy state of the ion⁷. However, a more general and efficient solution to the problem is to use phosphorescent emissive materials.

Phosphorescence is the forbidden relaxation of an excited state with spin symmetry different from the ground state; in organic molecules it typically results from a triplet to a singlet

relaxation. Although low-efficiency electroluminescence has been demonstrated⁸ at 100 K using the phosphorescent material benzo-phenone doped into poly(methylmethacrylate), no energy transfer from the host material was found. If a phosphorescent dye participates in energy transfer, then triplet behaviour within the host can be directly examined. This may enable the accurate determination of spin statistics and other triplet properties, as well as resulting in very high electroluminescent efficiency devices.

One relatively well studied red emitting phosphorescent dye is PtOEP. Porphine complexes are known to possess long-lived triplet states useful in oxygen detection⁹. The addition of platinum to the porphine ring reduces the phosphorescence lifetime by increasing spin-orbit coupling; the triplet states gain additional singlet character and vice versa. This also enhances the efficiency of intersystem crossing from the first singlet excited state to the triplet excited state. Transient absorption spectrometry gives a singlet lifetime in PtOEP of ~ 1 ps, and the fluorescence efficiency is extremely weak¹⁰. In contrast, the room-temperature phosphorescence efficiency of PtOEP in a polystyrene matrix is¹¹ 0.5 with an observed lifetime of 91 μ s.

Induced dipole (or Förster) energy transfer to the triplet state is disallowed by spin conservation. However, energy transfer may occur by the parallel combination of Förster transfer to the singlet state, along with electron exchange (Dexter energy transfer). Dexter transfer is the 'physical' coherent transfer of an exciton from a donor to an acceptor site¹² at a rate proportional to the orbital overlap of the donor and acceptor molecules. Consequently, it is a short-range process, attenuating exponentially with distance. The transfer rate is proportional to the spectral overlap of the two species because the donor exciton energy must closely match that of the acceptor. Dexter processes allow for the transfer of triplet excitons because only the total spin of the donor-acceptor complex is conserved. Thus it is possible for both singlets and triplets to excite phosphorescence, increasing the theoretical efficiency limit from that for fluorescence by a factor of four or larger.

PtOEP shows¹³ strong absorption at a wavelengths of 530 nm, corresponding to the peak emission of the electron transport material, tris-(8-hydroxyquinoline) aluminium (Alq₃); this makes PtOEP a suitable dopant for Alq₃-based organic light-emitting diodes (LEDs). To study the electroluminescent properties of Alq₃/PtOEP systems, organic LEDs were produced by high-vacuum (10^{-6} torr) thermal evaporation onto a cleaned glass substrate precoated with conductive, transparent indium tin oxide (ITO). The test structure consists of a 60-Å-thick layer of copper phthalocyanine to aid hole injection from the ITO, a 350-Å-thick layer of the hole transporting material 4,4'-bis[N-(1-naphthyl)-N-phenyl-amino] biphenyl (α -NPD), a 400-Å-thick layer of Alq₃ doped with varying molar concentrations of PtOEP, and a further 100-Å-thick layer of Alq₃ to prevent quenching of PtOEP excitons at the cathode. Finally, a shadow mask with 1-mm-diameter openings was used to define the cathode consisting of a 1,000-Å-thick layer of 25:1 Mg:Ag, with a 500-Å-thick Ag cap. All measurements were performed in air at room temperature except for the photoluminescence measurements which were performed under nitrogen.

The electroluminescence spectra of the devices with three different concentrations of PtOEP are shown in Fig. 1. No significant emission is found for the previously identified singlet state, expected at ~ 580 nm (ref. 10), but strong emission is observed from the triplet excited state at 650 nm, with weaker emission seen at the vibronic harmonic overtones at 623 nm, 687 nm and 720 nm. At high drive currents, the emission at 650 nm saturates, and increasing emission is seen from other features, especially the broad Alq₃ peak centred at 530 nm. This results in a reversible shift in colour from deep red to orange, corresponding to a shift in the device chromaticity co-ordinates shown in Fig. 1 inset.

Spectral and time-resolved emission measurements were performed with a streak camera on our organic LEDs excited with current pulses. Slow decay of between 10 and 50 μ s was observed for

‡ Permanent address: Department of Chemistry, Goucher College, Baltimore, Maryland 21204-2794, USA.

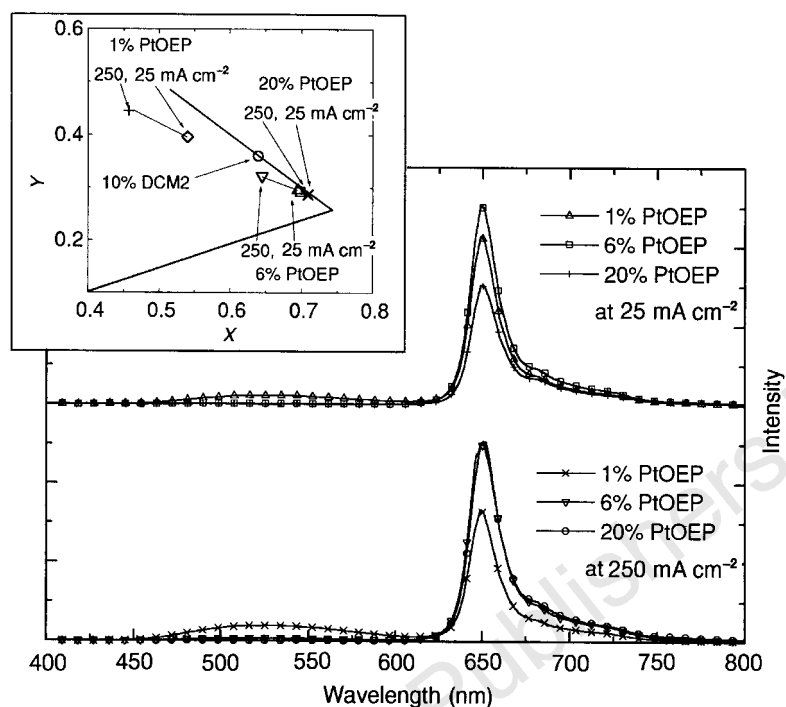


Figure 1 Spectra of the organic LEDs with different molar concentrations of PtOEP at different current densities. Upper traces, 1%, 6% and 20% PtOEP in Alq₃ organic LEDs at 25 mA cm⁻²; lower traces, 1%, 6% and 20% PtOEP in Alq₃ organic LEDs at 250 mA cm⁻². We note the increased Alq₃ emission at 530 nm in the 1% PtOEP organic LED. Insert, Commission Internationale de L'Éclairage (CIE)

chromaticity coordinates for the devices in the main figure at the specified current densities. Only the red corner of the CIE diagram is shown. We note the trend from saturated red to orange with increasing current. The 6% PtOEP in Alq₃ device at 25 mA cm⁻² has a luminance of 100 cd m⁻². The DCM2 result is taken from ref. 15.

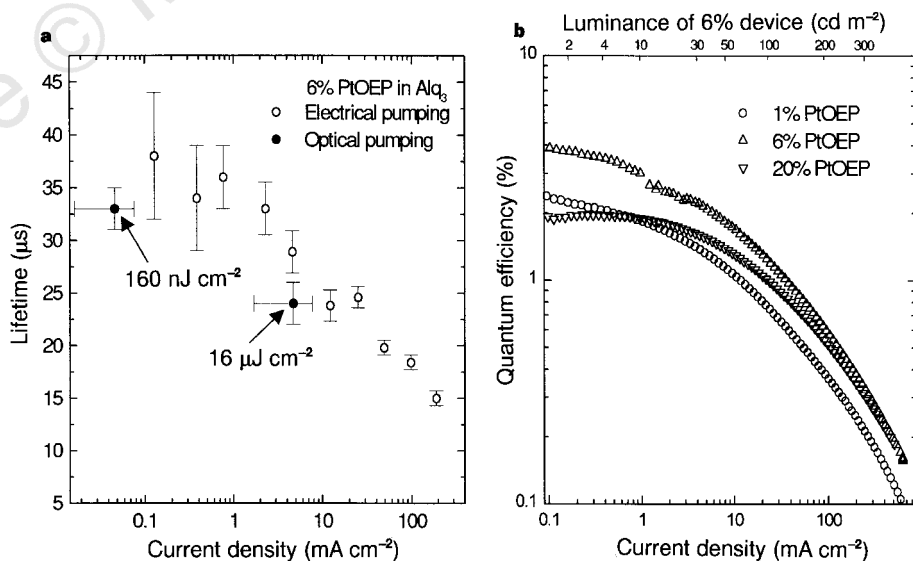


Figure 2 Lifetime and quantum efficiency of PtOEP emission. **a**, Phosphorescent lifetime of the 6% PtOEP in Alq₃ organic LED as function of current density (open circles). To explore the influence of bimolecular processes, PtOEP quenching in the absence of charge carriers was investigated using photoluminescence. This experiment was performed by pumping a 500-Å-thick film of 6% PtOEP in Alq₃ with 500-ps pulses from a nitrogen laser at a wavelength of 337 nm. At a pulse energy density of 16 μJ cm⁻², we measured a PtOEP lifetime of (24 ± 2) μs. This is similar to the lifetime obtained under electrical pumping at current densities of ~5 mA cm⁻². To check if the optical and electrical decay times have a similar origin, we note that the laser output is absorbed principally by Alq₃ with an

absorption coefficient²⁰ of ~3 × 10⁴ cm⁻¹. Hence, ~26% of the nitrogen laser pulse is absorbed, giving an exciton density of 1.4 × 10¹⁸ cm⁻³. Assuming 100% transfer of injected carriers to PtOEP excitons, the current density corresponding to this pumping level is shown above (filled circles). When the pulse energy is decreased to ~160 nJ cm⁻², the lifetime increases to (33 ± 2) μs. The trend in the lifetime of optically pumped samples, which has been previously attributed¹⁴ to bimolecular quenching, is consistent with the trend of the electroluminescent lifetime data, suggesting a common physical origin. **b**, Quantum efficiency of PtOEP emission as a function of doping concentration and current density. The top axis shows the luminance of the 6% PtOEP in Alq₃ device.

the triplet state, confirming the presence of phosphorescence. The luminescent lifetime was also found to decrease with increasing current density, as shown in Fig. 2a for a device with a 6% concentration of PtOEP in Alq₃.

From the spectral data, the fraction of output power in the red (subtracting the Alq₃ emission) can be determined as a function of current to provide the external quantum efficiency (Fig. 2b). As expected¹², the trend in phosphorescent lifetime matches the quantum efficiency results: the shorter lifetimes observed at high current densities corresponds to lower quantum efficiencies. The external quantum efficiency reaches a maximum value of 4% at a PtOEP concentration of 6%. To our knowledge, this is considerably higher than the highest external quantum efficiency obtained to date for red-emitting, fluorescent organic LEDs.

Because PtOEP molecules have a long exciton lifetime, it is possible that saturation is responsible for the decreased quantum efficiency at high currents and low dopant concentrations. This can be examined by comparing the total number of PtOEP sites to the maximum number of photons extracted. For a molar concentration of 1%, the number of PtOEP sites in the ~50-Å-thick recombination zone³ near the α -NPD/PtOEP:Alq₃ interface is $\sim 8 \times 10^{12} \text{ cm}^{-2}$. After adjusting for the output coupling efficiency⁶, the internal emitted photon flux saturates at $\sim 3 \times 10^{16} \text{ s}^{-1} \text{ cm}^{-2}$, yielding a radiative lifetime of $\sim 300 \mu\text{s}$, consistent with previous PtOEP phosphorescence efficiency and lifetime measurements¹¹. If sites within the recombination zone saturate, then Alq₃ excitons are less likely to transfer to PtOEP, corresponding to an increase in the probability for radiative recombination in Alq₃. This is supported by the data shown in Fig. 1, which shows increased Alq₃ emission for devices with low concentrations ($\sim 1\%$) of PtOEP.

Saturation of emissive sites is alleviated by increasing the concentration of PtOEP. Yet from Fig. 2b, it is evident that at molar concentrations $\geq 6\%$, the quantum efficiency again decreases. As high densities of porphyrin complexes often show bimolecular quenching¹⁴, it is likely that the PtOEP/Alq₃ system is also affected by this process. Hence, the decrease in quantum efficiency with increasing current could be partially accounted for by the decreasing

lifetime. As the current increases, the phosphorescent decay becomes bi-exponential, indicating the presence of a quenching process such as bimolecular recombination (see Fig. 2a legend).

The doping concentrations required to maximize quantum efficiency are extremely high compared to devices doped with fluorescent dyes and excited by long-range ($\sim 40 \text{ \AA}$) Förster transfer. This suggests that the shorter-range Dexter process may be dominant in the Alq₃:PtOEP system. Calculations of the extent of energy transfer from Alq₃ to PtOEP support this hypothesis. Accounting for photons lost via emission through the sides of the device and total internal reflection within the substrate⁶, a maximum external efficiency of 4% corresponds to an internal efficiency of 23%. PtOEP doped in polystyrene has a lifetime of 91 μs and a phosphorescent yield of 50% (ref. 11). Given a maximum exciton lifetime in the PtOEP/Alq₃ devices of 45 μs , we infer a peak phosphorescent yield of $\leq 25\%$. This implies $\geq 90\%$ energy transfer from Alq₃ to PtOEP, confirming that both triplets and singlet excitons must participate in energy transfer.

Conclusive evidence of Dexter transfer is obtained by examining the non-normalized spectra of the two devices shown in Fig. 3. A 100-Å-thick layer of Alq₃ doped with $\sim 1\%$ DCM2 is placed in the recombination zone of both devices. As DCM2 is fluorescent with efficient energy transfer from Alq₃ (ref. 15), this layer effectively removes singlet excitons from the devices. Remaining singlets eventually recombine in Alq₃, yielding the small shoulder in the spectra at $\sim 530 \text{ nm}$. However, in device 2, an additional layer of $\sim 10\%$ PtOEP in Alq₃ is introduced 200 Å away from the recombination zone. In this device, emission is seen from PtOEP without any change in the intensity of emission from either DCM2 or Alq₃. Hence, PtOEP cannot be an efficient trap, as carriers removed by PtOEP in device 2 would result in a decrease in the DCM2 and Alq₃ emission; an effect clearly not observed. As the DCM2 acts as a filter that removes singlet Alq₃ excitons, the only possible origin of the PtOEP luminescence is Alq₃ triplet states that have diffused through the DCM2 and intervening Alq₃ layers.

Given the observed decrease of quantum efficiency with increasing current, it is important that Alq₃ devices doped with PtOEP achieve a useful luminosity before significant quenching occurs. Unlike some red dyes¹⁶ with peak emission around 650 nm, PtOEP emission is saturated and does not extend significantly into the infrared, thereby maximizing the luminosity. At a molar concentration of 6%, a quantum efficiency of 1.3% and a power efficiency of 0.15 lm W^{-1} at 100 cd m^{-2} is obtained. In comparison with other compounds with saturated red emission, such as the Eu complexes^{17–19}, PtOEP possesses quantum and power efficiencies that are superior by at least an order of magnitude. Compounds with unsaturated red emission, such as DCM2 and its variants¹⁶, possess comparable quantum efficiencies but emit closer to orange than PtOEP; further red-shifting by increasing the DCM2 concentration causes the quantum efficiency to decrease, along with a considerable increase in infrared emission¹⁵.

Although PtOEP demonstrates the efficiency improvements made possible by the participation of triplet excitons in phosphorescence, the long lifetime of PtOEP ($>10 \mu\text{s}$) and the short-range nature of Dexter energy transfer cause saturation of emission at low dopant concentrations. It is also observed that bimolecular interactions quench emission at high exciton densities, causing a decrease in quantum efficiency for high currents and doping concentrations. But it is possible that other host materials could be used with PtOEP to increase efficiency even beyond the high values reported here. Given the performance of PtOEP, other phosphorescent dyes emitting in the blue and green regions of the spectrum also present an attractive area of study for display applications. □

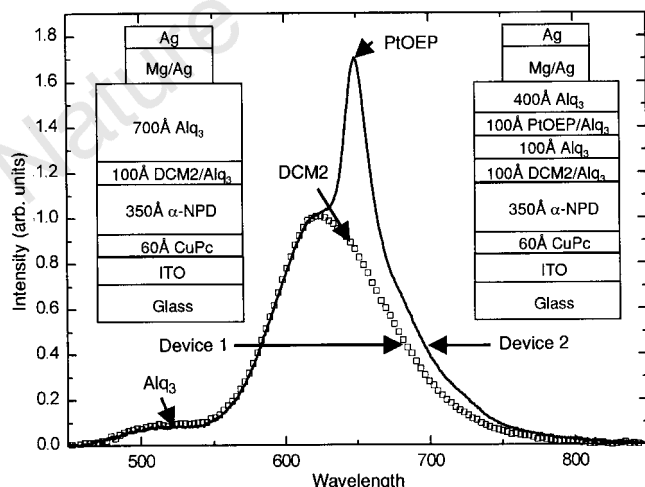


Figure 3 Two electroluminescent devices demonstrating that Alq₃ triplets are transferred to PtOEP. Each device contains a 100-Å-thick layer of $\sim 1\%$ DCM2 in Alq₃ at the recombination zone. This layer acts to remove singlet states. Remaining singlets recombine in Alq₃, yielding the shoulder apparent in the spectra at 530 nm. Device 2 contains an additional layer of $\sim 10\%$ PtOEP in Alq₃ positioned 200 Å away from the Alq₃/ α -NPD interface. Strong emission is seen from the PtOEP without a corresponding decrease in emission from DCM2 or Alq₃. The spectra were measured at a current density of 6 mA cm^{-2} . (DCM2 is [2-methyl-6-[2-(2,3,6,7-tetrahydro-1H,5H-benzo[*ij*]quinolin-9-yl) ethenyl]-4H-pyran-4-ylidene] propane-dinitrile.)

Received 20 April; accepted 24 June 1998.

1. VanSlyke, S. A. & Tang, C. W. US Patent No. 4539507 (1985).
2. Burroughes, J. H. *et al.* Light emitting diodes based on conjugated polymers. *Nature* **347**, 539–541 (1990).

3. Tang, C. W., VanSlyke, S. A. & Chen, C. H. Electroluminescence of doped organic thin films. *J. Appl. Phys.* **65**, 3610–3616 (1989).
4. Tsutsui, T. & Saito, S. *Organic Multilayer-Dye Electroluminescent Diodes—Is There Any Difference with Polymer LED?* (Kluwer Academic, Dordrecht, 1993).
5. Rothberg, L. J. & Lovinger, A. J. Status of and prospects for organic electroluminescence. *J. Mater. Res.* **11**, 3174–3187 (1996).
6. Gu, G. *et al.* High-external-quantum-efficiency organic light-emitting devices. *Opt. Lett.* **22**, 396–398 (1997).
7. Dirr, S. *et al.* Vacuum-deposited thin films of lanthanide complexes: spectral properties and applications in organic light-emitting devices. In *SID 97 Digest* (ed. Morreale, J.) 778–781 (Soc. for Information Display, Santa Ana, CA, 1997).
8. Hoshino, S. & Suzuki, H. Electroluminescence from triplet excited states of benzophenone. *Appl. Phys. Lett.* **69**, 224–226 (1996).
9. Mills, A. & Lepre, A. Controlling the response characteristics of luminescent porphyrin plastic film sensors for oxygen. *Anal. Chem.* **69**, 4653–4659 (1997).
10. Ponterini, G., Serpone, N., Bergkamp, M. A. & Netzel, T. L. Comparison of radiationless decay processes in osmium and platinum porphyrins. *J. Am. Chem. Soc.* **105**, 4639–4645 (1983).
11. Papkovski, D. B. New oxygen sensors and their applications to biosensing. *Sens. Actuators B* **29**, 213–218 (1995).
12. Turro, N. J. *Modern Molecular Photochemistry* (University Science Books, Mill Valley, CA, 1991).
13. Liu, H.-Y., Switalski, S. C., Coltrain, B. K. & Merkel, P. B. Oxygen permeability of sol-gel coatings. *Appl. Spectrosc.* **46**, 1266–1272 (1992).
14. Rodriguez, J., McDowell, L. & Holten, D. Elucidation of the role of metal to ring charge-transfer states in the deactivation of photo-excited ruthenium porphyrin carbonyl complexes. *Chem. Phys. Lett.* **147**, 235–240 (1988).
15. Bulovic, V. *et al.* Bright, saturated, red-to-yellow organic light-emitting devices based on polarization-induced spectral shifts. *Chem. Phys. Lett.* **287**, 455–460 (1998).
16. Chen, C. H., Tang, C. W., Shi, J. & Klubeck, K. P. Improved red dopants for organic electroluminescent devices. *Macromol. Symp.* **125**, 49–58 (1997).
17. Tsutsui, T., Takada, N., Saito, S. & Ogino, E. Sharply directed emission in organic electroluminescent diodes with an optical-microcavity structure. *Appl. Phys. Lett.* **65**, 1868–1870 (1994).
18. Kido, J. *et al.* Bright red light emitting organic electroluminescent devices having a europium complex as an emitter. *Appl. Phys. Lett.* **65**, 2124–2126 (1994).
19. Sano, T. *et al.* Novel europium complexes for electroluminescent devices with sharp red emission. *Jpn. J. Appl. Phys.* **34**, 1883–1887 (1994).
20. Garbuzov, D. Z., Bulovic, V., Burrows, P. E. & Forrest, S. R. Photoluminescence efficiency and absorption of aluminum-tris-quinolate (Alq₃) thin films. *Chem. Phys. Lett.* **249**, 433–437 (1996).

Acknowledgements. We thank V. G. Kozlov for help with the transient measurements, and P. E. Burrows for discussions. This work was supported by Universal Display Corporation, DARPA, AFPSR and NSF.

Correspondence and requests for materials should be addressed to S.R.F. (forrest@ee.princeton.edu).

Synthesis and organization of zeolite-like materials with three-dimensional helical pores

Thurman E. Gier, Xianhui Bu, Pingyun Feng & Galen D. Stucky

Department of Chemistry, University of California, Santa Barbara, California 93106, USA

The increasing demand for enantiomerically pure chemicals has stimulated extensive research into the preparation of heterogeneous chiral catalysts or separation media that combine both shape selectivity and enantioselectivity¹. Helical pores in inorganic materials might be able to perform such functions, but their occurrence is rare¹. Attempts have been reported to synthesize a specific enantiomorph of the chiral zeolite beta², and chiral metal complexes have been used to assemble inorganic precursors into chiral frameworks^{3,4}. Materials with a fully three-dimensional array of helical structural units are particularly rare, because helical structures (such as quartz) are commonly generated by a uni-dimensional symmetry element acting on an achiral structural subunit^{5,6}. Here we report on a family of zeolite-type materials (which we call UCSB-7) that possess two independent sets of three-dimensional crosslinked helical pores, separated by a gyroid periodic minimal surface¹³. We have synthesized the UCSB-7 framework for various compositions (zinc and beryllium arsenates, gallium germanate) using either inorganic cations or amines as structure-directing agents. The helical-ribbon motif that we identify might be exploited more widely for developing useful chiral solid-state structures.

The framework topology of UCSB-7 was determined from gallo-germanates synthesized with K⁺ or amines such as NH₂(CH₂)_nNH₂ (n = 2–5). We have grown large crystals of gallo-germanate phases with a variety of linear and branched mono-, di- or tri-amines

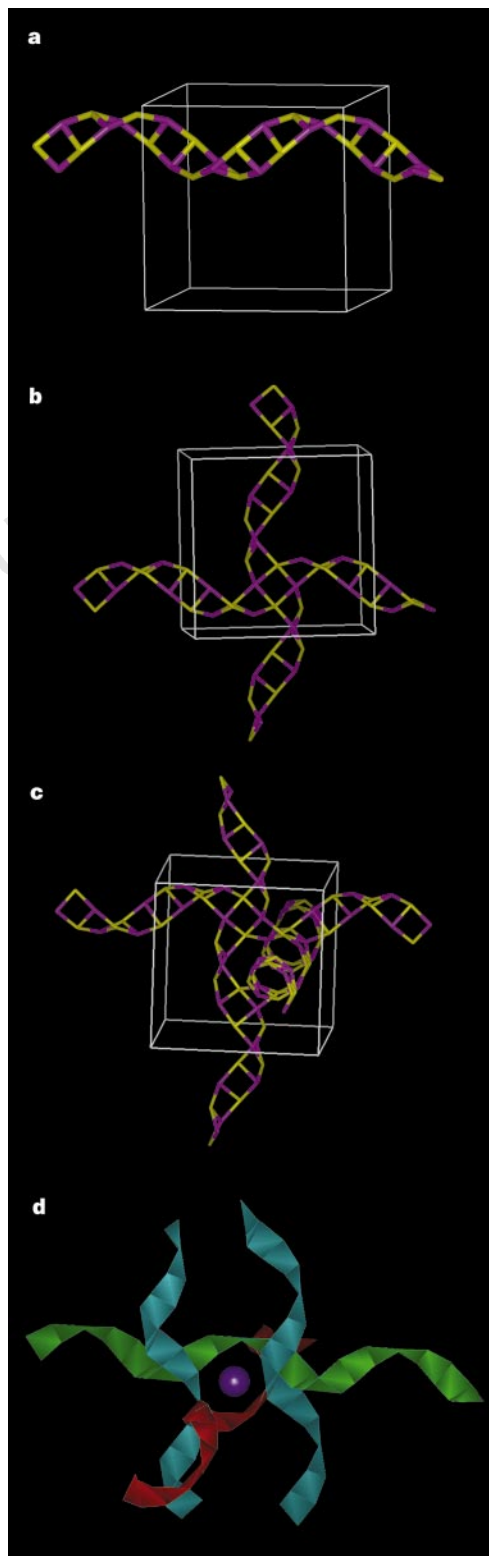


Figure 1 The helical ribbons are crosslinked in three dimensions to give 12-ring pores. From one to three dimensions (a–c): crosslinking of mutually perpendicular inorganic helical chains in UCSB-7K. In all panels the bridging oxygen atoms are omitted. The yellow sites represent Ge⁴⁺ atoms and the purple sites Ga³⁺ atoms. d, The K⁺ ion channel with the 12-ring pore. The purple sphere represents the K⁺ ion.

including methylamine, *N*-methylethylenediamine, tris(2-aminoethylamine), *N*-(2-aminoethyl)-1,3-propanediamine and 3,3'-diamino-*N*-methyldipropylamine. An exceptional case is the successful use of poly(ethylenimine), an organic polymer, in the synthesis of UCSB-7. We identified non-gallogermanate phases on the basis of powder diffraction patterns, all of which can be indexed by using a body-centred cubic cell with a repeat distance of ~ 18 Å. Another zeolite structure with the 18-Å body-centred cubic cell is ZK-5, but it has very different lattice symmetry ($Im\bar{3}m$), leading to a different powder diffraction pattern.

Of particular interest is the synthesis of the ethylenediamine zinc arsenate phase UCSB-7 at freezing temperatures (from -20 to 4°C). At 25°C , it transforms into UCSB-3 within a few days and at 100°C , UCSB-3 is formed within 20 min (ref. 7). The ethylenediamine gallogermanate phase is, however, much more stable and its crystals have been grown at 180°C , together with UCSB-3GaGe and GaGe-SOD1 (a triclinic sodalite analogue). The framework structure of UCSB-7 is not stable to water loss in the range 250 – 400°C depending on the chemical composition. However, essentially complete replacement of K^+ by NH_4^+ in saturated NH_4HCO_3 solution at 10°C was easily accomplished and established by the recovery of clean $\text{Ga}_2\text{Ge}_2\text{O}_7$ after pyrolysis at $1,000^\circ\text{C}$. The Ge/Ga

ratio in gallogermanates can be manipulated by using different structure-directing agents. For example, in UCSB-7dma (here, dma is dimethylamine), the Ge/Ga ratio is 3 on the basis of crystal structure analysis and electron-probe elemental analysis.

UCSB-7 consists of two helical-T-O-T-atomic chains (T refers to tetrahedral atoms) with the same handedness that are coupled together to form a helical ribbon made up of edge-sharing 4-rings (Fig. 1). The pitch of the helix is equal to the unit cell repeat length and the width is ~ 5.7 Å for gallogermanate phases based on the distance between T-atoms. Each helical turn contains eight 4-rings (Fig. 1). The cylindrical volume within each helical ribbon forms a channel along a unit cell axis, and, when projected along the axis, the channel seems to have an 8-ring aperture (see Fig. 2a). In fact there is no closed 8-ring anywhere in the structure: only 4-, 6- and 12-rings are present. There are too few atoms on the helical ribbons to completely isolate the channel spaces along the axes of the helices. However, the connectivity of the helical ribbons (see Figs 1 and 3) orientated along the three cubic axes does give rise to 12-ring windows and the pore structure (see Figs 2 and 3). When viewed down the cubic unit cell body-diagonal directions, the disposition of these 12-ring channels is easily seen (Fig. 2b).

The interconnection by 12-ring windows (Fig. 3a) between the

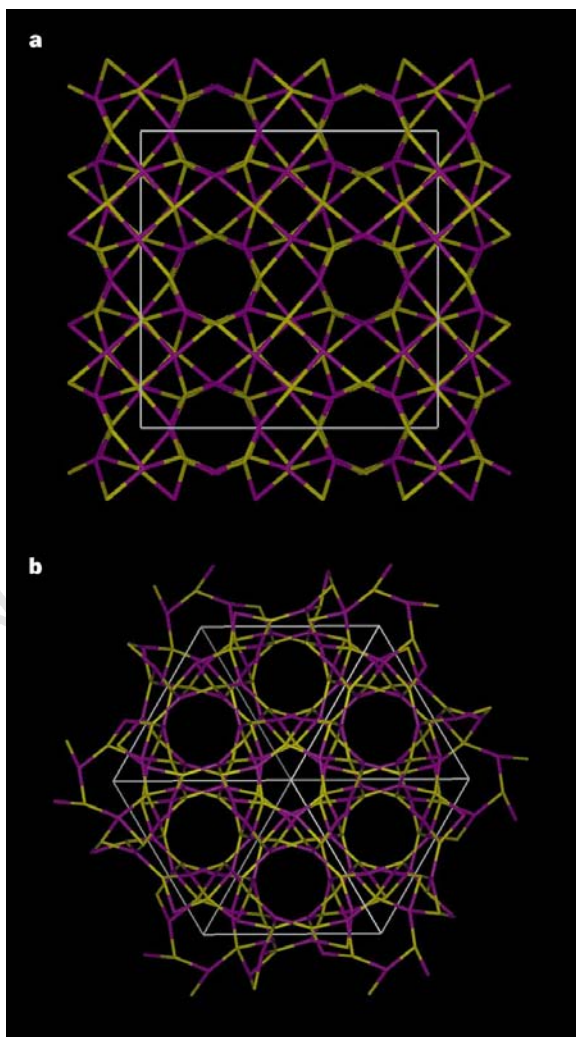


Figure 2 Helices viewed parallel to cubic unit cell edges and body diagonal directions. **a**, The framework structure of UCSB-7K viewed down the *a* axis showing channels that look like 8-ring pores in projection. The apparent walls between channels are in fact other channels that are perpendicular to channels shown here. **b**, The framework structure of UCSB-7K viewed down the crystallographic [111] direction.

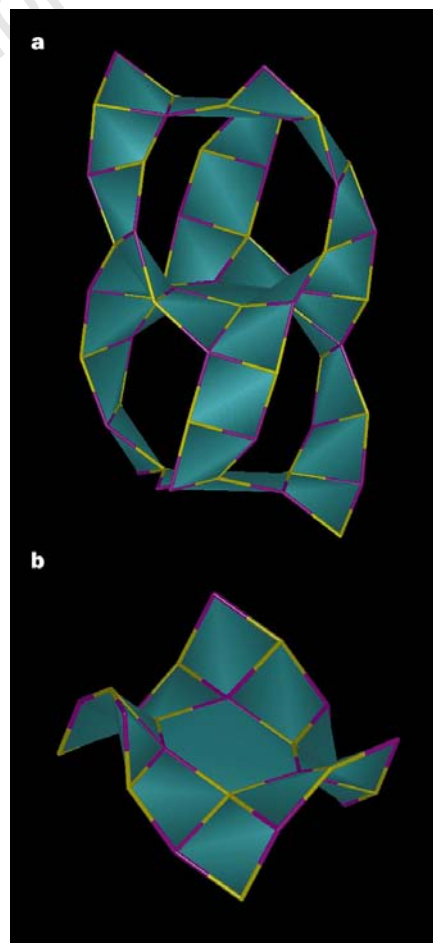


Figure 3 The crystalline gyroid minimal surface of UCSB-7 separates the space into two disconnected volumes. **a**, Two sets of 12-ring windows (three above and three below) are shown here. They connect helices of opposite handedness into two equal but separate volumes. In the middle, a monkey saddle detailed in **b** separates the top portion from the lower one. **b**, A section of the framework surface that can be described as the monkey saddle typical of a gyroid minimal surface. The flat point coincides with the centre of the 6-ring.

void space within helical ribbons occurs only between helices of the same handedness. Every other helical ribbon in Fig. 2a viewed parallel to the cell edge is of opposite handedness, and in general for the structure there are an equal number of right-handed and left-handed helices. Thus, for each helical ribbon, there are four surrounding helical ribbons of the opposite handedness ($a/2\text{ Å}$ between channel centres; a is the cell length) and four others of the same handedness ($a\sqrt{2}/2\text{ Å}$ between channel centres) that are propagating along the same crystallographic direction (Fig. 2a). The interconnections between helices of the same handedness form two equal and disconnected volumes, which in zeolite terms are two independent sets of three-dimensional 12-ring intersecting channels with opposing handednesses. Such a structural feature has not previously been observed in a zeolite-type structure⁸. In fact, even materials with one set of three-dimensional 12-ring channels are rare among known zeolite structures, with zeolite beta, faujasite, hexagonal faujasite (EMC-2), UCSB-6 and UCSB-10 being the only examples^{9,10}.

The wall that separates the right-handed and left-handed regions can be described as a gyroid bicontinuous minimal surface that has also been found in liquid crystal phases and mesoporous MCM-48¹¹, except in this case the wall structure is crystalline and well defined at the atomic level. Illustrated in Fig. 3b is a portion of the UCSB-7 wall structure representing a monkey saddle, which is a typical feature of a gyroid surface¹². For UCSB-7, flat points of such a minimal surface coincide with centres of all 6-rings located at unit cell origins, body-, face- and edge-centres, and on body-diagonals one quarter distance ($a\sqrt{3}/4\text{ Å}$) from cell origins. Another zeolite-type structure that has been described by a gyroid surface is a relatively dense phase, analcime¹³.

Because both right- and left-handed helices are present, the crystal space group symmetry would be centric $Ia\bar{3}d$ if the tetrahedral framework atoms were all the same or disordered. However, in the case of UCSB-7K, the best results were obtained with an alternative distribution of Ga and Ge sites when the refinement was performed in a non-centric, chiral space group $I2_13$. This was confirmed by the observed difference in metal–oxygen bond distances, which are 1.838 Å for Ga–O and 1.762 Å for Ge–O, in close agreement with those in a gallium germanate sodalite analogue that has bond distances of 1.839 Å for Ga–O and 1.745 Å for Ge–O¹⁴. With this ordering, the right- and left-handed helices are not mirror images and thus the overall chiral symmetry of UCSB-7 structures is related to framework chemical compositions and ordering of T-atom sites. For example, in UCSB-7 the magnitude of the chirality is determined by the difference in the electrostatic potential at the gallium (negative) and germanium (neutral) atom framework sites, whereas in the zinc or beryllium arsenate structures the chirality would be related to the difference in the electrostatic potentials associated with the zinc or beryllium (negative) and the arsenic (positive) atom framework sites.

We suggest that the formation of the helical pattern is related to the presence of some geometrical features observed in UCSB-7 structures. UCSB-7K has a relatively small T–O–T angle (minimum, maximum and average are 122.3°, 127.0° and 124.2° respectively), consistent with the T–O–T angle found in the gallium germanate sodalite analogue (125.5°) and the T–O–T angles in two arsenate sodalite structures, $\text{Na}_3(\text{ZnAsO}_4)_3 \cdot 4\text{H}_2\text{O}$ and $\text{Li}_4\text{Cl}(\text{BeAsO}_4)_3$ of 123.8° and 123.4°, respectively^{14,15}. Structurally the small T–O–T angles can be accommodated in a 4-ring that has a UDUD (where U is up and D is down) configuration of the T atoms and therefore the building-block surface curvature and geometry required to generate the helical configuration.

Although a significant amount of the current research on zeolite synthesis is directed at low-charge framework structures, UCSB-7 suggests that there is still a great potential for generating novel zeolite-type structures in the framework charge domain represented by traditional aluminium-rich zeolites¹⁶. □

Methods

(KGaGeO₄)₆·7H₂O (UCSB-7K). To a polypropylene bottle were added 9.26 g 2 M K₃GaO₄ (12 mmol), 15 ml H₂O and 17.47 g 1 M K₆GeO₅ (13 mmol). To the resulting clear solution was added 10.26 g of concentrated HNO₃ (114 mmol), producing a slowly developing gel slurry. After 1 day at 100 °C, the well-settled crystalline powder was filtered off, washed and recovered (2.61 g, 82% yield). The powder diffraction pattern could be indexed by using a body-centred cubic cell. Crystals suitable for single crystal X-ray diffraction were grown by holding the gel slurry at 100 °C for three months (the crystal growth of amine-directed phases usually takes one week at 180 °C). Thermogravimetric analysis to 600 °C gave 8.0% weight loss (calc. = 7.9%). Crystallographic data for (KGaGeO₄)₆·7H₂O are: space group $I2_13$; $a = 18.6766(1)\text{ Å}$; $V = 6514.69(6)\text{ Å}^3$, $Z = 8$, $\text{MoK}\alpha$, $2\theta_{\text{max}} = 50^\circ$, $R(F) = 3.82\%$ for 103 parameters and 1,920 reflections with $I > 2\sigma(I)$.

(NaGaGeO₄)₆·xH₂O (UCSB-7Na). To a polypropylene bottle was added 8.80 g 2 M Na₃GaO₄ (12 mmol), 16.04 g 1 M Na₄GeO₄ (13 mmol) and 15 ml H₂O. To the clear solution was then added 8.91 g concentrated HNO₃ (99 mmol), giving a thick milky precipitate. After 7 h at 100 °C, the precipitate was well settled and was recovered (2.945 g, 98% yield). The powder diffraction pattern could be indexed by using a body-centred cubic cell ($a = 18.608\text{ Å}$). Thermogravimetric analysis showed 9.3% weight loss (calc. 8.4%), leaving only NaGaGeO₄ at 950 °C.

(NaBeAsO₄)₆·xH₂O (x ≈ 7, UCSB-7Be). 4.68 g Na₂HAsO₄·7H₂O (15 mmol), 2.142 g 4 M NaOH (7.46 mmol) and 10 ml H₂O were added to a polypropylene bottle to give a clear basic solution; 6.98 g 2 M Be(NO₃)₂ (12 mmol) were then added to give a gel slurry with a pH of ~6. After being left for ~20 h at 100 °C, the slurry was well settled and the pH had dropped to 4. Standard filtration and washing techniques yielded 2.17 g (91% yield) of a white powder whose diffraction pattern could be indexed by using a body-centred cubic cell ($a = 17.272\text{ Å}$). Thermogravimetric analysis showed 11.0% weight loss to 800 °C (calc. 10.9%).

(KZnAsO₄)₆·xH₂O (x ≈ 7, UCSB-7Zn). 18.60 g 4.8 M H₃AsO₄ (62.75 mmol), 31.03 g 2 M Zn(NO₃)₂ (48 mmol) and 20 ml H₂O were added to a polypropylene bottle and then 38.04 g of 5 M KOH was added to the clear solution. The thick gel slurry produced had a pH of ~8. After being left overnight at 25 °C (UCSB-7Zn can also be made at 4 °C, but is unstable at 100 °C), the slurry was well settled and on filtration and washing yielded 11.8 g (91% yield) of a white microcrystalline powder whose diffraction pattern could be indexed by using a body-centred cubic cell ($a = 18.710\text{ Å}$). Thermogravimetric analysis showed a 9.5% weight loss to 450 °C (calc. = 7.9% for $x = 7$) occurring in two steps. The residue was determined as being monoclinic KZnAsO₄ by X-ray diffraction. Use of the same procedure, but with up to 20% of the zinc replaced by Co²⁺, Mn²⁺ or Fe²⁺, resulted in clean single-phase products (blue, faint pink and green–grey, respectively), indicating the possible replacement of zinc sites by transition metals. The maximum possible substitution levels have not been determined.

(EDA)₃(ZnAsO₄)₆·xH₂O (UCSB-7EDA). 2.40 g ethylenediamine (40 mmol), 10 ml H₂O and 7.71 g 2 M Zn(NO₃)₂ (12 mmol) were added to a small polypropylene bottle. The initial precipitate redissolved and the clear solution was then cooled to ~4 °C in a refrigerator. After the solution had chilled, 4.74 g 4 M H₃AsO₄ (14 mmol) was added and the resulting gel slurry was returned to the refrigerator (pH 8.5). After 2 days at 4 °C, the gel was well crystallized and recovery yielded 1.15 g (33% yield) of material whose powder pattern was indexed by using a body-centred cubic cell ($a = 18.957\text{ Å}$). UCSB-7EDA can also be made at –20 °C at a slower crystallization speed.

Received 27 April; accepted 13 July 1998.

1. Baiker, A. Chiral catalysis on solids. *Curr. Opin. Solid State Mater. Sci.* **3**, 86–93 (1998).
2. Davis, M. E. New vistas in zeolite and molecular sieve catalysis. *Acc. Chem. Res.* **26**, 111–115 (1993).
3. Bruce, D. A., Wilkinson, A. P., White, M. G. & Bertrand, J. A. The synthesis and characterization of an aluminophosphate with chiral layers: trans-Co(dien)₂·Al₃P₄O₁₆·3H₂O. *J. Solid State Chem.* **125**, 228–233 (1996).
4. Morgan, K., Gainsofrid, G. & Milestone, N. A novel layered aluminum phosphate [Co(en)Al₃P₄O₁₆·3H₂O] assembled about a chiral metal complex. *J. Chem. Soc. Chem. Commun.* 425–426 (1995).
5. Harrison, W. T. A., Gier, T. E., Stucky, G. D., Broach, R. W. & Bedard, R. A. NaZnPO₄·H₂O, an open-framework sodium zincophosphate with a new chiral tetrahedral framework topology. *Chem. Mater.* **8**, 145–151 (1996).
6. Soghomonian, V., Chen, Q., Haushalter, R. C., Zubiet, J. & O'Connor, C. J. An inorganic double helix: hydrothermal synthesis, structure, and magnetism of chiral [(CH₃)₂NH₂][K₄(V₁₀O₁₀(H₂O)₂(OH)₄(PO₃)₄)]·4H₂O. *Science* **259**, 1596–1599 (1993).

7. Bu, X., Feng, P., Gier, T. E. & Stucky, G. D. Two ethylenediamine templated zeolite type structure zinc arsenate and cobalt phosphate systems. *J. Solid State Chem.* **136**, 210–215 (1998).
8. Smith, J. V. Topochemistry of zeolites and related materials. 1. Topology and geometry. *Chem. Rev.* **88**, 148–182 (1988).
9. Meier, W. M., Olson, D. H. & Baerlocher, C. *Atlas of Zeolite Structure Types* (Elsevier, Boston, 1996).
10. Bu, X., Feng, P. & Stucky, G. D. Large-cage zeolite structures with multidimensional 12-ring channels. *Science* **278**, 2080–2085 (1997).
11. Monnier, A. *et al.* Cooperative formation of inorganic–organic interfaces in the synthesis of silicate mesostructures. *Science* **261**, 1299–1303 (1994).
12. Andersson, S. & Jacob, M. *The Mathematics of Structures* (R. Oldenbourg, München, 1997).
13. Andersson, S., Hyde, S. T., Larsson, K. & Lidin, S. Minimal surfaces and structures: from inorganic and metal crystals to cell membranes and biopolymers. *Chem. Rev.* **88**, 221–242 (1988).
14. Nenoff, T. M. *et al.* Structural and chemical investigations of $\text{Na}_3(\text{ABO}_4)_3 \cdot 4\text{H}_2\text{O}$ -type sodalite phases. *Inorg. Chem.* **33**, 2472–2480 (1994).
15. Harrison, W. T. A., Gier, T. E. & Stucky, G. D. Two lithium chloroberyllo(phosphate/arsenate) sodalites: $\text{Li}_2\text{Cl}(\text{BePO}_4)_3$ and $\text{Li}_2\text{Cl}(\text{BeAsO}_4)_3$. *Acta Crystallogr. C* **50**, 471–473 (1994).
16. Feng, P., Bu, X. & Stucky, G. D. Hydrothermal synthesis and structural characterization of zeolite analogue compounds based on cobalt phosphate. *Nature* **388**, 735–741 (1997).

Supplementary Information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. This research was supported in part by the National Science Foundation and the Office of Naval Research. P.F. thanks the Materials Research Laboratory for a Corning Foundation Fellowship. All figures were generated using software programs from Molecular Simulations.

Correspondence and requests for materials should be addressed to G.D.S. (e-mail: stucky@chem.ucsb.edu).

Heterogeneous production of nitrous acid on soot in polluted air masses

M. Ammann*, M. Kalberer*†, D. T. Jost*, L. Tobler*, E. Rössler*, D. Piguet*, H. W. Gäggeler*† & U. Baltensperger*

* Paul Scherrer Institute, Laboratory of Radio- and Environmental Chemistry, CH-5232 Villigen, Switzerland

† University of Bern, Department of Chemistry and Biochemistry, CH-3012 Bern, Switzerland

Polluted air masses are characterized by high concentrations of oxidized nitrogen compounds which are involved in photochemical smog and ozone formation. The OH radical is a key species in these oxidation processes. The photolysis of nitrous acid (HNO_2), in the morning, leads to the direct formation of the OH radical and may therefore contribute significantly to the initiation of the daytime photochemistry in the polluted planetary boundary layer. But the formation of nitrous acid remains poorly understood: experimental studies imply that a suggested heterogeneous formation process involving NO_2 is not efficient enough to explain the observed night-time build-up of HNO_2 in polluted air masses¹. Here we describe kinetic investigations which indicate that the heterogeneous production of HNO_2 from NO_2 on suspended soot particles proceeds 10^5 to 10^7 times faster than on previously studied surfaces. We therefore propose that the interaction between NO_2 and soot particles may account for the high concentrations of HNO_2 in air masses where combustion sources contribute to air pollution by soot and NO_x emissions. We believe that the observed HNO_2 formation results from the reduction of NO_2 in the presence of water by C–O and C–H groups in the soot. Although prolonged exposure to oxidizing agents in the atmosphere is likely to affect the chemical activity of these groups, our observations nevertheless suggest that fresh soot may have a considerable effect on the chemical reactions occurring in polluted air.

There is strong evidence that the spatial and temporal distribution of HNO_2 throughout the planetary boundary layer (PBL) is governed by continuous formation and daytime photolysis, which result in pronounced diurnal cycles with night-time maxima of up to more than 10 parts per billion by volume (p.p.b.v.). Correlation

analysis points to NO_2 as a source component, and typical NO_2 conversion rates inferred are 10^{-6} s^{-1} (refs 2–4). The gas phase reaction of OH radicals with NO may only contribute significantly to the very low HNO_2 levels found during the day. HNO_2 emitted directly from combustion processes, together with other oxidized nitrogen compounds and soot, is of minor importance⁵. Therefore, a heterogeneous process involving NO_2 has been suggested. The maximum surface area per unit volume, S/V, attributed to ground and airborne surfaces is 10^{-2} cm^{-1} and 10^{-5} cm^{-1} , respectively, for typical mixing heights (minimum, 100 m during the night) in the PBL. These surfaces have been proposed to catalyse the reaction between NO_2 and H_2O to form nitric and nitrous acid (reaction (1)), but measured HNO_2 formation rates using different types of bulk surfaces (solid materials and aqueous solutions) were found to be two to five orders of magnitude smaller than the NO_2 to HNO_2 conversion rates inferred from atmospheric observations¹.

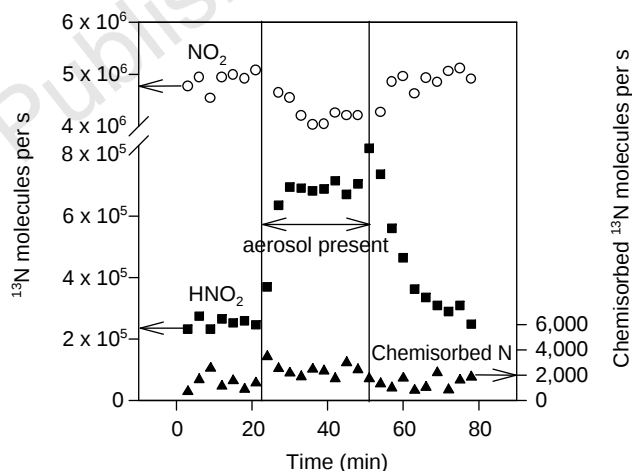
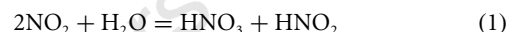


Figure 1 Online record of the rate of H^{13}NO_2 formation from 12 p.p.b.v. NO_2 in the absence and presence of soot particles. The data, obtained over a period of 80 min, record the rate of H^{13}NO_2 formation 20 s after two different gas flows have been mixed in a flow reactor at 22 °C, ambient pressure and 50% relative humidity. The first gas flow, containing NO_2 (enriched with $^{13}\text{NO}_2$) in He and 20% O_2 , is introduced at a rate of $5 \text{ cm}^3 \text{ s}^{-1}$. The $^{13}\text{NO}_2$ content is adjusted so that the $^{13}\text{NO}_2$ molecules are fed into the reactor at a rate of $5 \times 10^6 \text{ s}^{-1}$. The NO_x -free soot aerosol is produced with a pulsed spark discharge between graphite rods in Ar, followed by dilution with air and a thermodenuder operating at 400 °C. O_2 , N_2 and H_2O impurities in the Ar (99.998%) give rise to soot-like complex organic carbon compounds on the particle surfaces. Before introducing the gas mixture containing the aerosol into the reactor at a rate of $13.3 \text{ cm}^3 \text{ s}^{-1}$, part of the flow is passed through a differential mobility analyser coupled to a condensation particle counter to monitor the concentration and size distribution of the aerosol. The total flow rate through the 80-cm-long glass reactor (2.5 cm inner diameter) is $18.3 \text{ cm}^3 \text{ s}^{-1}$, and the gas mixture contains 14% O_2 , 47% N_2 , 13% He, 26% Ar, 1.2% H_2O and 12 p.p.b.v. NO_2 . For product detection and identification, 5 compound-specific 40-cm-long glass denuders (4 mm inner diameter) were designed as coils around the γ -ray detectors (5.5 cm diameter). The first was coated with sodium chloride (HNO_3 detection), the second and third with a sodium carbonate/glycerol mixture (HNO_2), the fourth with a triethanolamine/guaiacol mixture (NO_2), and the fifth with cobalt(II/III) oxide (NO). For HNO_2 , the two denuders with the same coating in series allowed a correction for interference by NO_2 . The γ -ray activities measured at the denuders (NO_2 and HNO_2) and at the filter (N_{chem}) were converted to fluxes of molecules into each trap using the 9.96-min half-life of ^{13}N . In each experiment, the background effect caused by the reaction of NO_2 to HNO_2 at the glass walls of the flow reactor (reaction (1)) was measured before and after the soot particles were present.

Significant chemical effects of soot from both natural and anthropogenic sources in the atmosphere have been postulated⁶, but previous experimental studies on the reaction of NO₂ with soot to form NO or HNO₂ were mostly performed with bulk carbonaceous samples in vacuum or with very high NO₂ concentrations (refs 7–9). That is, the reaction has not yet been studied under realistic conditions with respect to the physical and chemical characteristics of the soot surface, pressure, NO₂ concentration and humidity. The present study addresses this shortcoming by making use of the short-lived radioactive isotope ¹³N (half-life 9.96 min) as a tracer^{10,11} to monitor the effects of soot aerosol on NO and HNO₂ formation under conditions typical for the polluted troposphere.

The interaction between NO₂ and soot particles was studied in a flow reactor at 22 °C, ambient pressure and 50% relative humidity. A flow of inert gas and oxygen containing 12 p.p.b.v. NO₂ (enriched with ¹³NO₂) and a flow of an air/inert-gas mixture containing suspended soot particles, both equilibrated to a relative humidity of 50%, were mixed in the reaction vessel. The extent of the interaction between NO₂ and soot aerosol was monitored by combining ¹³N measurements with the denuder technique^{11,12} which differentiates between chemical compounds according to their reactivity. This set-up allows the simultaneous and online measurement of nitric acid (HNO₃), HNO₂, NO₂, NO and particle-bound products (chemisorbed nitrogen, N_{chem}) in the gas leaving the reaction vessel.

The soot aerosol was monitored online by differential mobility analysis¹³ giving concentrations of typically 2×10^6 particles cm⁻³ with an average diameter of 70 nm and S/V of 3×10^{-4} cm⁻¹. This size distribution agrees with measured black carbon size distributions in the polluted PBL¹⁴. For non-spherical particles, the equivalent mobility diameter, which is the diameter of the sphere with the same mobility, is the preferred size parameter when mass transfer to agglomerated aerosol particles is considered¹⁵. The total surface area density derived with this technique for weighed samples of particles collected on a filter was 300 m² g⁻¹. Carbon and oxygen chemistry on the surface of the soot particles was checked by X-ray photoelectron spectroscopy¹⁶ which is sensitive to chemical bonding of the elements present in the uppermost layers of the surface. The spectra in the C(1s) and O(2p) region were similar to samples directly collected from diesel exhaust, indicating that the soot particles used in the present experiments expose similar C–O and C–H functionalities as fresh diesel soot particles. Thermal analysis¹⁷ showed that they contained 40% organic and 60% elemental carbon which again is typical for soot from combustion exhaust¹⁸. The soot particles used here thus seem to adequately represent the chemical and physical characteristics of fresh soot emitted to the atmosphere by combustion sources, while still being suitable for a very sensitive laboratory aerosol study.

Figure 1 shows the fluxes of HNO₂, NO₂ and N_{chem} that are detected 20 s after the gas flows containing the reagents have mixed in the reaction vessel. In the absence of soot aerosol, the NO₂, HNO₂ and N_{chem} fluxes are fairly constant, with only relatively small amounts of HNO₂ and N_{chem} being formed. In the presence of soot aerosol, however, the HNO₂ flux increases significantly, thus providing direct evidence for substantial heterogeneous HNO₂ production. To obtain an estimate for the rate of production, HNO₂ concentrations were inferred from the measured fluxes and monitored as a function of reaction time (the time that has elapsed after the reagent flows have mixed in the reaction vessel). Figure 2 shows data obtained in the presence of 12 p.p.b.v. NO₂ and for reaction times ranging from 5 to 155 s. The data indicate that HNO₂ is produced under these conditions at a rate of 1.0×10^{12} molecules per s per cm² aerosol surface. The data also imply that HNO₂ formation is much more rapid within the first 5 s after the reaction has been initiated. With a modified experimental set-up that improved the time-resolution of the system, a HNO₂ production rate of 3.3×10^{13} molecules per s per cm² aerosol surface was

derived on this time scale (Fig. 3). The probability that a collision between a NO₂ molecule and a soot particle will result in the formation of HNO₂ is known as the reaction probability γ . The data shown in Figs 2 and 3 yield values of 1.1×10^{-2} and 3.3×10^{-4} , respectively, for this parameter. The reaction probability of reaction (1) at 50% relative humidity, in contrast, is only 3×10^{-9} as inferred from laboratory studies on other surfaces (ref. 1), thus indicating that the interaction between NO₂ and the soot aerosols studied here leads to a HNO₂ formation rate that is about 5 to 7 orders of magnitude faster.

The previously suggested heterogeneous process (reaction (1)) (ref. 19) predicts the formation of equal amounts of HNO₂ and HNO₃, but we were not able to detect the latter as a product of the NO₂/soot interaction. Although it is conceivable that HNO₃ remains adsorbed to the particle surface, or that it decomposes into NO (ref. 8), these processes are not likely to explain the observed absence of significant HNO₃ formation: the concentration of particle-bound N-species we measured was two orders of magnitude lower than the amount of HNO₂ formed while NO formation was below the detection limit (1% of the NO₂ concentration) of our system. These observations imply that, instead of the disproportionation of N(IV) into equal amounts of N(V) and N(III) species, a net reduction of N(IV) to N(III) is taking place at low NO₂ concentrations. Although previous studies^{8,9,19} have already indicated that NO₂ may be reduced on soot to form NO or N₂O, our observations suggest the existence of an additional reaction channel (reaction (2)) that results in HNO₂ formation.

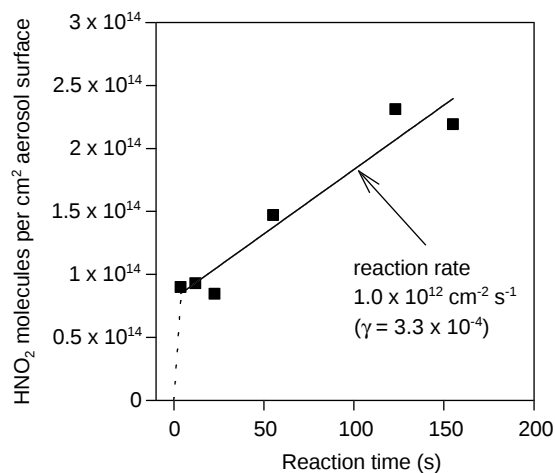
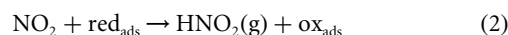


Figure 2 HNO₂ concentrations, normalized to total particle surface area, as a function of reaction time. The data were obtained using the system described in Fig. 1. The reaction times could be varied between 5 and 20 s by using a movable inlet for the introduction of the NO₂ gas flow, thereby achieving a range of distances (and hence reaction times) between the mixing of the two gas flows and the detection of reaction products. The longer reaction times of 50 s, 114 s and 155 s were achieved by using 3 larger reaction vessels with fixed inlets and diameters of 4 cm, 6 cm and 7 cm. After subtraction of the background signals (obtained as in Fig. 1) for each experiment, the HNO₂ concentration was calculated from the net flux of ¹³N molecules into the denuder and the gas flow rate of 16.7 cm³ s⁻¹ through the denuders, and assuming constant ratios of labelled to non-labelled molecules for all species. HNO₂ formation is normalized to the total particle surface area calculated by integration of the size spectra obtained with the differential mobility analyser during each experiment. The reaction probability, γ , was calculated from the linear regression of the data (solid line) and the gas kinetic flux of NO₂ molecules to the particle surface. The dashed line indicates the fast onset of HNO₂ production at times below 5 s.

Reaction (2) is presented as a net reaction without explicitly stating its elementary steps, where red_{ads} and ox_{ads} denote a reduced and an oxidized site on the soot surface, respectively. The electron transfer properties of the carbon-oxygen functionalities typical for soot⁷ are not sufficiently understood to allow a specific assignment of the different kinetic regimes found (Figs 2 and 3). Our results clearly indicate that HNO_2 forms much more efficiently than NO. The slow NO formation was quantified in our earlier aerosol study¹¹, and is in contrast to reports on bulk amorphous carbon samples giving a reaction probability of 10^{-1} for NO formation^{8,9}.

We performed several tests to explore the sensitivity of reaction (2) to oxygen, ozone and humidity. Exposure of the soot particles to NO_x -free air of 50% relative humidity for 1 min before the reaction with NO_2 did not change the HNO_2 output, suggesting that oxygen does not compete with NO_2 for reducing groups on the soot surface. A corresponding pretreatment of the particles with 200 p.p.b.v. ozone (O_3) for 2 min reduced HNO_2 formation by 40%. On the other hand, HNO_2 formation was not affected when the particles were exposed to 12 p.p.b.v. NO_2 and up to 200 p.p.b.v. O_3 at the same time, indicating that ozone will not significantly influence HNO_2 formation on soot in the atmosphere in the short term. Whereas reaction (1) was found to be very sensitive to the water concentration¹⁹, in our experiments, at 0.5% relative humidity the formation of HNO_2 was (within the experimental uncertainty) the same as at 60%. Even at these low humidities, the gas-phase concentration of H_2O is large enough to allow some soot functionalities (such as carboxyl groups) to be hydrated or to lead to the presence of physisorbed H_2O (ref. 8).

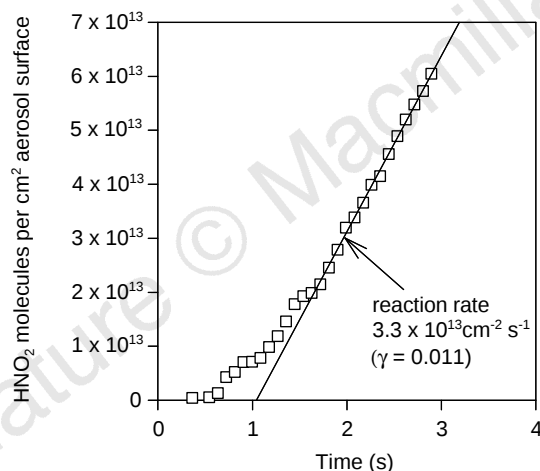


Figure 3 HNO_2 formation on soot particles at 12 p.p.b.v. NO_2 and 50% relative humidity at reaction times below 3 s. The aerosol and the NO_2 flows were mixed at the inlet of a 1-m-long tube (0.8 cm inner diameter) coated with a sodium carbonate/glycerol mixture which retains all HNO_2 molecules diffusing to the tube wall. The total flow rate was $16.7 \text{ cm}^3 \text{ s}^{-1}$. The amount of ^{13}N arriving at each tube segment of 3 cm length was measured with a γ -ray detector after 30 min. First, the background signal caused by NO_2 flowing through the tube was measured without the soot particles in the gas stream. For the plot, the difference of the signals with and without the aerosol at each segment was integrated along the tube and normalized to the particle surface area in the gas stream. The first 30 cm were not used for the kinetic analysis, because mixing of the two gas flows occurs here. Under the conditions chosen, the diffusion to the wall was faster than production of new HNO_2 molecules, so that no correction was applied to the data. The diffusion causes a slight delay of the response at the wall (corresponding to 0.5 s on the time scale) which does not affect the slope analysed for the calculation of the reaction probability, γ . This experiment differs from the experiments shown in Figs 1 and 2 in that HNO_2 is continuously removed from the gas phase during the NO_2 /aerosol interaction. The data shown in Figs 1 and 2, in contrast, illustrate the formation rates that are obtained when HNO_2 is removed at the end of a fixed reaction time.

Urban air masses typically contain¹⁴ $10 \mu\text{g m}^{-3}$ soot, which corresponds to an aerosol surface to volume ratio of $3 \times 10^{-5} \text{ cm}^{-1}$ (assuming a specific surface area identical to that measured for our model aerosol). Using these values and the reaction probabilities obtained from our experiments, we calculate night-time NO_2 conversion rates of $3.7 \times 10^{-3} \text{ s}^{-1}$ and $1.1 \times 10^{-4} \text{ s}^{-1}$ at 10 p.p.b.v. NO_2 and 50% relative humidity for the two different kinetic regimes. The long-term kinetics of reaction (2) could not be studied due to the low half-life of ^{13}N , but already within the short timescale of our study (155 s), the observed HNO_2 production ($3 \times 10^{14} \text{ molecules cm}^{-2}$) builds up 0.3 p.p.b.v. HNO_2 . Extrapolated to longer timescales (10 h night-time) or lower soot loadings (0.1 to $1 \mu\text{g m}^{-3}$ for rural areas), concentrations typical for the PBL are reached (0.01 to 10 p.p.b.v.; ref. 1). After high HNO_2 levels have built up during the night, the yield of OH radicals from its photolysis in the morning (maximum, $3 \times 10^7 \text{ cm}^{-3} \text{ s}^{-1}$ for 1 p.p.b.v. HNO_2) exceeds other photochemical sources of OH radical which reach their maximum around noon (typically $10^7 \text{ cm}^{-3} \text{ s}^{-1}$)⁴. Hence, the high reactivity of soot particles in the polluted PBL may be one of the main factors in initiating daytime photochemistry. The suggested HNO_2 production on soot particles may also play an important role in less polluted rural areas where a fast HNO_2 production process has been postulated to bring measured HNO_2 values and a photochemical model into agreement²¹.

Soot particles have been suggested^{6,22} to affect the nitrogen oxide chemistry of the remote troposphere and stratosphere by reducing HNO_3 and NO_2 to NO. We find, in contrast, that the reduction of NO_2 in polluted air masses leads to at least two orders of magnitude more HNO_2 than NO. Furthermore, the HNO_2 that has been built up throughout the night is photolysed the next morning to form the strongly oxidizing OH radical, and it therefore seems unlikely that heterogeneous reactions on soot particles result in a net reduction of the nitrogen oxide species in polluted air. The process may even have a net oxidizing effect on nitrogen oxides, and hence on ozone in the PBL, but in the absence of detailed model calculations a reliable assessment of its overall effect is not yet possible.

Soot particles are ubiquitous throughout the atmosphere²³, including the upper troposphere and lower stratosphere²⁴. During their long residence times in the atmosphere they undergo a variety of ageing processes, such as the proposed reaction with NO_2 , reactions with other oxidants, and coating with sulphuric acid or secondary organic material. On aged soot, the HNO_2 formation process suggested here may only be significant if surface species oxidized by reaction (2) are recycled. SO_2 could be a good candidate to maintain such a mechanism²⁵, and solar radiation could affect the fate of the oxidized surface by inducing decarboxylation²⁶ or other photon-induced formation of new surface radicals²⁷. Although HNO_2 formation on soot in remote areas of the PBL has already been suggested (ref. 28) to mediate the nitration of hydrocarbon trace species, a reliable assessment of the overall effects of this process requires further studies of the heterogeneous chemistry of aged soot and its associated secondary constituents. □

Received 18 July 1997; accepted 23 June 1998.

- Lammel, G. & Cape, J. N. Nitrous acid and nitrite in the atmosphere. *Chem. Soc. Rev.* 361–369 (1996).
- Kitto, A.-M. N. & Harrison, R. M. Nitrous and nitric acid measurements at sites in south-east England. *Atmos. Environ.* 26, 235–241 (1992).
- Andrés-Hernández, M. D., Notholt, J., Hjorth, J. & Schrems, O. A DOAS study on the origin of nitrous acid at urban and non-urban sites. *Atmos. Environ.* 30, 175–180 (1996).
- Harrison, R. M., Peak, J. D. & Collins, G. M. Tropospheric cycle of nitrous acid. *J. Geophys. Res.* 101, 14429–14439 (1996).
- Kirchstetter, T. W., Harley, R. A. & Littlejohn, D. Measurement of nitrous acid in motor vehicle exhaust. *Environ. Sci. Technol.* 30, 2843–2849 (1996).
- Lary, D. J. et al. Carbon aerosols and atmospheric photochemistry. *J. Geophys. Res.* 102, 3671–3682 (1997).
- Akhter, M. S., Chughtai, A. R. & Smith, D. M. The structure of hexane soot I: Spectroscopic studies. *Appl. Spectrosc.* 39, 143–153 (1985).
- Rogaski, C. A., Golden, D. M. & Williams, L. R. Reactive uptake and hydration experiments on amorphous carbon treated with NO_2 , SO_2 , O_3 , HNO_3 , and H_2SO_4 . *Geophys. Res. Lett.* 24, 381–384 (1997).
- Tabor, K., Gutzwiller, L. & Rossi, M. J. Heterogeneous chemical kinetics of NO_2 on amorphous carbon at ambient temperature. *J. Phys. Chem.* 98, 6172–6175 (1994).

10. Baltensperger, U. *et al.* Use of positron-emitting ^{13}N for studies of the selective reduction of NO by NH_3 over vanadia/titania catalyst at very low reactant concentrations. *J. Phys. Chem.* **97**, 12325–12330 (1993).
11. Kalberer, M. *et al.* Heterogeneous chemical processing of $^{13}\text{NO}_2$ by monodisperse carbon aerosols at very low concentrations. *J. Phys. Chem.* **100**, 15487–15493 (1996).
12. Allegrini, I. *et al.* Annular denuder method for sampling reactive gases and aerosols in the atmosphere. *Sci. Tot. Environ.* **67**, 1–16 (1987).
13. Liu, B. Y. H. & Pui, D. Y. H. A submicron aerosol standard and the primary absolute calibration of the condensation nuclei counter. *J. Colloid Interface Sci.* **47**, 155–171 (1974).
14. Berner, A. *et al.* Modal character of atmospheric black carbon size distributions. *J. Geophys. Res.* **101**, 19559–19565 (1996).
15. Rogak, S. N., Baltensperger, U. & Flagan, R. C. Measurement of mass transfer to agglomerate aerosols. *Aerosol Sci. Technol.* **14**, 447–458 (1991).
16. Wild, U., Pfänder, N. & Schlögl, R. Species analysis of automotive carbon particles: application of XPS for integral analysis of filter samples. *Fresenius J. Anal. Chem.* **357**, 420–428 (1997).
17. Cachier, H., Bremond, M.-P. & Buat-Ménard, P. Determination of atmospheric soot carbon with a simple thermal method. *Tellus B* **41**, 379–390 (1989).
18. Bertscher, H., Künzel, S., Hügl, Ch. Characterization of particles in combustion engine exhaust. *J. Aerosol Sci.* **29**, 389–396 (1998).
19. Chughtai, A. R., Welch, J. W. E., Akhter, M. S. & Smith, D. M. Spectroscopic study of gaseous products of soot—oxides of nitrogen/water reactions. *Appl. Spectrosc.* **44**, 294–298 (1990).
20. Febo, A. & Perrino, C. Prediction and experimental evidence for high air concentration of nitrous acid in indoor environments. *Atmos. Environ.* **A 25**, 1055–1061 (1991).
21. Staffelbach, T., Neftel, A. & Horowitz, L. W. Photochemical oxidant formation over southern Switzerland 2. Model results. *J. Geophys. Res.* **102**, 23363–23373 (1997).
22. Hauglustaine, D. A., Ridley, B. A., Solomon, S., Hess, P. G. & Madronich, S. HNO_3/NO_x ratio in the remote troposphere during MLOPEX2: Evidence for nitric acid reduction on carbonaceous aerosols? *Geophys. Res. Lett.* **23**, 2609–2612 (1996).
23. Lioussé, C. *et al.* A global three-dimensional model study of carbonaceous aerosols. *J. Geophys. Res.* **101**, 19411–19432 (1996).
24. Blake, D. F. & Kato, K. Latitudinal distribution of black carbon soot in the upper troposphere and lower stratosphere. *J. Geophys. Res.* **100**, 7195–7202 (1995).
25. De Santis, F. & Allegrini, I. Heterogeneous reactions of SO_2 and NO_2 on carbonaceous surfaces. *Atmos. Environ.* **26**, 3061–3064 (1992).
26. Smith, D. M. & Chughtai, A. R. Photochemical effects in the heterogeneous reaction of soot with ozone at low concentrations. *J. Atmos. Chem.* **26**, 77–91 (1997).
27. Weschler, C. J., Mandich, M. L. & Graedel, T. E. Speciation, photosensitivity, and reactions of transition metal ions in atmospheric droplets. *J. Geophys. Res.* **91**, 5189–5204 (1986).
28. Ciccioli, P., Cecinato, A., Brancaleoni, E., Frattoni, M. & Zaccari, P. Formation and transport of 2-nitrofluoranthene and 2-nitropyrene of photochemical origin in the troposphere. *J. Geophys. Res.* **101**, 19567–19580 (1996).

Acknowledgements. The stable proton beam from the Philips Cyclotron necessary for ^{13}N production was provided by the staff of the Accelerator Facilities at Paul Scherrer Institute; V. Lavanchy, B. Schnyder and R. Koetz performed the offline analysis of soot samples. Contributions from S. Nyeki, C. Zellweger, E. Weingartner, P. Zimmermann and B. Eichler were greatly appreciated.

Correspondence and requests for materials should be addressed to M.A. (e-mail: markus.ammann@psi.ch).

Mantle discontinuities and temperature under the North American continental keel

Aibing Li*, Karen M. Fischer*, Michael E. Wyssession† & Timothy J. Clarke‡

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

† Department of Earth and Planetary Sciences, Washington University, St Louis, Missouri 63130, USA

‡ Department of Earth and Environmental Sciences, New Mexico Institute of Mining and Technology, Socorro, New Mexico 87801, USA

A ubiquitous feature of upper-mantle seismic velocity models has been the presence of high-velocity ‘keels’ beneath stable continental interiors^{1–5}. Uncertainty remains, however, regarding the maximum depth to which continental keels extend, the degree to which they have cooled the mantle that surrounds them and their role in mantle flow. Here we investigate thermal anomalies across the eastern margin of the North American continental keel by imaging the seismic discontinuities at depths of 410 and 660 km with compressional-to-shear converted waves recorded by a 1,500-km-long seismometer deployment in the eastern United States. The thickness of the transition zone (the region nominally between depths of 410 and 660 km) and the depth to the ‘410-km’ discontinuity indicate that cold keel material and sub-keel downwellings must be largely confined to the upper mantle and may

impinge on the transition zone only in localized regions and with thermal anomalies of less than ~ 150 K. A 20-km depression of the ‘660-km’ discontinuity to the south of the westernmost stations coincides with a region of fast velocity in the deep transition zone² and may be associated with the remnants of the subducted Farallon plate^{1,2,4}.

Discontinuities in seismic velocity have been observed at depths of roughly 410 km and 660 km in the Earth’s mantle (hereafter referred to as the ‘410’ and ‘660’). These discontinuities mark the upper and lower boundaries of the mantle transition zone, and are generally interpreted as mineral phase transitions from α -olivine to β -spinel, and from γ -spinel to perovskite+magnesiowüstite, respectively. The Clapeyron slopes of the equilibrium phase boundaries indicate that the depth of the ‘410’ discontinuity should be deflected upwards in regions of colder temperature, whereas the ‘660’ should be deflected downwards. The topography of the ‘410’ and ‘660’ discontinuities can therefore provide information about thermal structure in the mantle transition zone.

Knowledge of transition-zone temperatures below continental regions is important in understanding the role of continental keels in mantle flow. High-velocity keels have been imaged beneath stable continental interiors to depths of 300 km or more, but although keel velocity anomalies are strongest in the upper mantle, their maximum depth extent has not been clearly resolved^{1–7}. Some studies have argued that continental keels are neutrally buoyant features in which positive density anomalies due to cold temperature are roughly balanced by negative density anomalies due to chemical depletion^{5,8,9}. However, other investigations have concluded that keels are in fact anomalously dense, and should be associated with broad zones of strong mantle downwelling^{10,11}. In addition, even if keels are neutrally buoyant, the presence of these cold anomalies in the mantle over timescales of 1–2 Gyr may produce small-scale convective instabilities at the base of the keel¹². Transition-zone discontinuity topography can constrain the maximum depth extent of the cold keel itself, and if downwelling is occurring, the strength and wavelength of these convective instabilities.

Seismological investigations have reached a variety of conclusions regarding transition-zone discontinuity topography across continental

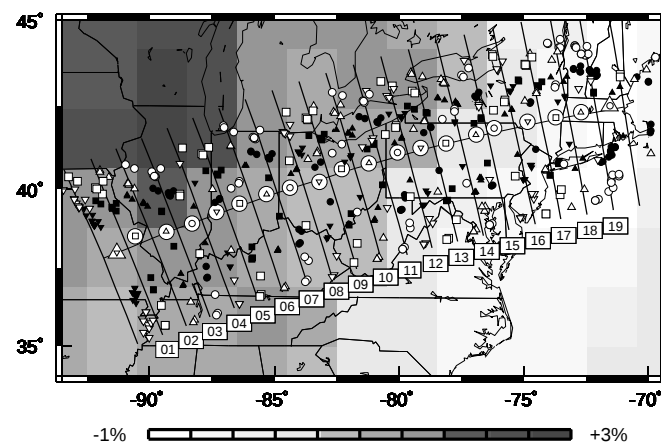


Figure 1 Ps conversion points at depths of 410 km (small black symbols) and 660 km (small white symbols) plotted on S-wave velocity anomalies from Grand² for depths of 175–250 km. Temporary stations of the MOMA array are shown as larger white circles, permanent stations HRV and CCM are denoted by larger white triangles, and the small symbol plotted at each station corresponds to the symbol of nearby Ps conversions recorded by that station. Bin numbers are given below the centre of each bin. The margin of the continental keel is marked by the west-to-east decrease in velocity anomaly, and is well sampled by Ps conversions from depths near 410 and 660 km.

interiors. Analyses of P-wave to S-wave (Ps) conversions from the '410' and '660' recorded at widely spaced stations across North America and Eurasia^{13–15} found very little variation in transition-zone thickness between orogenic and cratonic regions, whereas studies with denser sampling in localized regions reveal a thinning of the transition zone¹⁶ or a deepening of the '410'¹⁷ from young orogenic belts to stable continental interiors. ScS reverberations¹⁸ and some global studies of SS precursors⁶ find no evidence for consistent '410' depth or transition-zone thickness variations between continents and oceans, but other SS precursor studies argue for a shallower '410'¹⁷ or thicker transition zone¹⁹ beneath continental shields.

The goal of this study is to image transition-zone discontinuities beneath the eastern margin of the North American continental keel using Ps conversions recorded by the 1995–96 Missouri to Massachusetts Broadband Seismometer Experiment (MOMA; Fig. 1). Fast velocities that characterize the continental keel extend to depths of more than 300 km beneath the western stations of the array, and the edge of the keel is marked by a decrease in velocity in the 100 to 300 km depth range beneath the stations in Pennsylvania (longitude $\sim 80^\circ\text{W}$; Fig. 1). Beneath the easternmost stations in New England, the fast-velocity lithosphere appears to be much thinner (100 km or less)^{1–3}. The eastern margin of the North American keel has experienced little tectonic activity since the Palaeozoic Appalachian orogenies and the Triassic rifting associated with the opening of the Atlantic Ocean.

To determine the depths of the '410' and '660' discontinuities

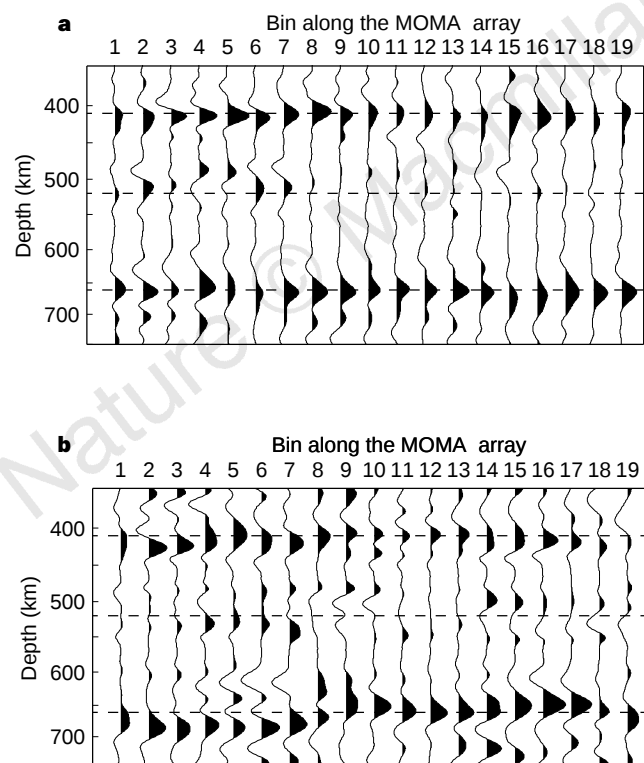


Figure 2 Ps receiver function stacks from bins to the north (a) and to the south (b) of the MOMA array. Each trace contains three segments (350–450 km, 450–550 km and 550–750 km) which were corrected for moveout and lateral heterogeneity at assumed conversion depths of 410 km, 520 km and 660 km, respectively. Corrections for lateral heterogeneity in the mantle and crust were obtained from the Grand² model, assuming a $\delta \ln(v_s)/\delta \ln(v_p)$ value of 1.0, and the crustal model that best fits crustal conversions recorded at the given station. Clear positive-polarity Ps conversions from depths near 410 and 660 km are imaged along the entire MOMA array, but a '520' phase is not consistently observed. The larger number of Ps phases in the bins to the north of the array yields receiver function stacks with a higher signal/noise ratio.

across the keel margin, we analysed the timing of Ps conversions generated at these discontinuities relative to the direct P phase. To enhance the amplitude of the Ps signals, receiver functions were calculated for individual seismograms by deconvolving vertical from radial seismogram components, corrected for moveout and lateral heterogeneity in the crust and the mantle, and stacked as a function of Ps conversion point location (Fig. 1). Corrections for moveout (variations in P–Ps time with distance) were calculated using the IASP91 radial velocity model²⁰. Variations in crustal structure across the array were estimated by modelling Moho Ps conversions and reverberations with sedimentary rock and crustal

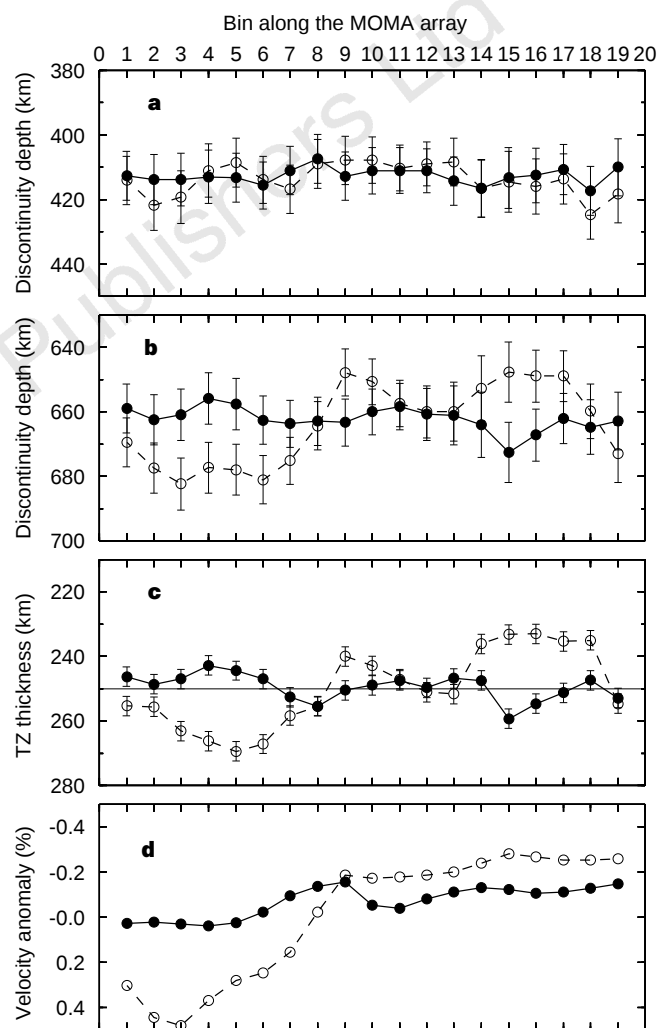


Figure 3 Discontinuity topography, transition zone (TZ) thickness, and deep transition-zone velocity anomalies beneath the MOMA array. **a**, Depth of the '410' discontinuity obtained from Ps receiver function stacks in the northern bins (filled circles) and southern bins (open circles). **b**, Depth of the '660' discontinuity in the northern bins (filled circles) and southern bins (open circles). **c**, Transition-zone thickness in the northern bins (filled circles) and southern bins (open circles). Relative Ps–P times between stations were obtained by cross-correlating stacked waveforms. Error bars correspond to the uncertainty in corrections to Ps–P times for crust and mantle heterogeneity. **d**, Average S-wave velocity anomalies from the Grand² model at 525–650 km depth in the northern bins (filled circles) and southern bins (open circles). The '410' discontinuity is relatively flat in both northern and southern bins, and in the northern bins, transition-zone thickness shows little variation, evidence that the broad, strong thermal anomalies associated with the continental keel do not reach transition-zone depths. In the southern bins, the transition zone thickens, the '660' deepens and velocities become faster east-to-west along the array, suggesting that cold temperatures exist in the deep transition zone to the south of the western stations.

basement layers and reasonable v_p and v_p/v_s values^{21,22} (here v_p and v_s are respectively P-wave and S-wave velocity). Lateral heterogeneity corrections for the crust and mantle were obtained for individual phases by ray-tracing through the three dimensional S-wave model of either Grand *et al.*² (hereafter referred to as the Grand model) or van der Lee and Nolet³, altered to include the best-fitting crustal model for a given station. P-wave velocities were obtained assuming v_p/v_s ratios from IASP91²⁰ and $\delta \ln(v_s)/\delta \ln(v_p)$ ratios of 1.0 and 2.5 (refs 2, 4, 23, 24).

Ps receiver function stacks for a given hypothetical discontinuity depth were calculated by combining Ps phases generated in common areas of the discontinuity surface. Phases with conversion points to the north and to the south of the stations were stacked separately, using 19 overlapping bins that span the MOMA array from west to east (Fig. 1). Receiver function stacks after corrections for moveout and lateral heterogeneity (obtained assuming the Grand² model and a $\delta \ln(v_s)/\delta \ln(v_p)$ value of 1.0) are shown in Fig. 2. Clear positive-polarity Ps conversions from depths near 410 and 660 km are imaged along the entire MOMA array. Although a few bins contain weak conversions from depths near 520 km, a '520' phase is not consistently observed. The average depths of the '410' km and '660' km discontinuities are 413 ± 3 km and 662 ± 4 km for the northern bins and 414 ± 5 km and 664 ± 12 km for the southern bins.

The sensitivity of the apparent discontinuity depths to assumed mantle velocity structure was evaluated by comparing receiver function stacks for the Grand² and van der Lee and Nolet³ S-wave models, assuming $\delta \ln(v_s)/\delta \ln(v_p)$ ratios of 1.0 and 2.5. These models predict relative discontinuity topographies along the array that differ by 4 km or less. Increasing the $\delta \ln(v_s)/\delta \ln(v_p)$ ratio from 1.0 to 2.5 deepens the discontinuities beneath the western stations by as much as 4 km. Because both models describe mantle heterogeneity beneath the MOMA array in terms of isotropic velocity, the effects of anisotropic velocity structure were estimated separately using shear-wave splitting in SKS phases recorded at the MOMA stations²⁵. Synthetic P and Ps phases (calculated using a propagator matrix method for a variety of anisotropic upper-mantle models that match the ~ 1.2 s of observed SKS splitting) indicate that the maximum possible difference in Ps–P time due to anisotropy in our receiver function stacks is 0.3 s, and that net effect of anisotropy on the '410' and '660' discontinuities is less than 3 km. Uncertainties in crustal v_p and v_p/v_s ratios^{21,22} correspond to ± 1.5 km of discontinuity topography beneath the western stations, and less than ± 1 km beneath the eastern stations. Combining the uncertainties in crust and mantle heterogeneity, we estimated error bars for the depths of the '410' and '660' discontinuities in each bin (Fig. 3a, b). Because transition-zone thicknesses inferred from P_{660s} – P_{410s} times are insensitive to upper-mantle velocity structure, and because transition-zone velocity heterogeneity and anisotropy are relatively weak^{1–4,26}, the error in transition-zone thickness is estimated to be 3 km or less (Fig. 3c).

If large-scale cold temperatures associated with the North American continental keel impinged on the '410' discontinuity, this discontinuity would be elevated and transition-zone thickness would be greatest in the most western bins to the north of the array which penetrate the furthest towards the centre of the keel (Fig. 1). However, transition-zone thickness shows little significant variation in the bins to the north of the array (Fig. 3c), and the '410' discontinuity (Fig. 3a) is relatively flat along the array in both the northern and southern bins. These results indicate that strong large-scale keel-related thermal anomalies must be confined to depths of less than 410 km. Weak variations in temperature and/or composition may occur at scale-lengths of 500 km or less, although they are not required by the data. Assuming Clapeyron slopes of 2.7 to 2.9 MPa K^{–1} for the olivine to β -spinel phase change^{27,28}, any such thermal anomalies must be less than 150 K in amplitude.

In contrast to the northern bins which show a roughly uniform transition-zone thickness of 250 ± 5 km, transition-zone thickness

in the southern bins does vary significantly from anomalously thin values beneath the eastern stations to anomalously thick values beneath the western stations (Fig. 3c). However, this variation in transition-zone thickness is due almost entirely to topography on the '660' discontinuity (Fig. 3b). Velocity anomalies in the Grand² model at depths of 525–650 km, although weak, are well correlated with the apparent '660' topography (Fig. 3d). Where velocities in this layer are fast, the '660' is deep, suggesting that the '660' depression in the southern bins at the western end of the array is thermal in origin. Assuming Clapeyron slopes of -2.1 to -3.0 MPa K^{–1} for the γ -spinel to perovskite+magnesiowüstite phase change^{27,29}, 20 km of topography on the '660' discontinuity implies a temperature anomaly of -270 K to -380 K. If real, these cold temperatures could be related to the subducted Farallon plate velocity anomaly that extends from depths of 650 km beneath the western MOMA stations to the core–mantle boundary in the Grand model^{1,2}, and in an independent P-wave travel-time inversion⁴. However, in the van der Lee and Nolet³ Rayleigh wave inversion, the Farallon plate anomaly appears much farther to the west in the transition zone³⁰, and attribution of the '660' depression to cold temperatures in the subducted Farallon plate remains speculative.

The lack of '410' discontinuity topography correlated with the North American keel margin beneath the MOMA array argues that the keel does not produce strong long-wavelength cold anomalies in the transition zone. This result is consistent with Ps phases recorded in cratonic and orogenic regions of Canada which reveal transition zone thickness variations of ± 5 km or less¹³. These images of transition-zone discontinuity topography provide fundamental constraints on geodynamical models that seek to explain the stability of continental keels over timescales of 1–2 Gyr. Although mantle downwelling that varies with time cannot be ruled out, cold North American keel material and related downwelling must at present be largely confined to the upper mantle. Cold keel material may impinge on the '410' discontinuity only in localized regions, if at all, and keel thermal anomalies in these areas must have amplitudes of 100 K or less. Similarly, any cold downwellings beneath the keel at transition-zone depths are very weak, small-scale, or both.

Received 23 December 1997; accepted 4 June 1998.

- Grand, S. P. Mantle shear structure beneath the Americas and surrounding oceans. *J. Geophys. Res.* **99**, 11591–11621 (1994).
- Grand, S. P., van der Hilst, R. D. & Widiyantoro, S. Global seismic tomography: a snapshot of convection in the earth. *GSA Today* **7**, 1–7 (1997).
- van der Lee, S. & Nolet, G. The upper mantle S-velocity structure of North America. *J. Geophys. Res.* **102**, 22815–22838 (1997).
- van der Hilst, R. D., Widiyantoro, S. & Engdahl, E. R. Evidence for deep mantle circulation from global tomography. *Nature* **386**, 578–584 (1997).
- Jordan, T. H. Structure and formation of the continental tectosphere. *J. Petrol.* (Special Lithosphere issue) 11–37 (1988).
- Flanagan, M. P. & Shearer, P. M. Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors. *J. Geophys. Res.* **103**, 2673–2692 (1998).
- Gu, Y., Dziewonski, A. M. & Agee, C. B. Global de-correlation of the topography of transition zone discontinuities. *Earth Planet. Sci. Lett.* **157**, 57–67 (1998).
- Forte, A. M., Dziewonski, A. M. & O'Connell, R. J. Continent-ocean chemical heterogeneity in the mantle based on seismic tomography. *Science* **268**, 386–388 (1995).
- Doin, M.-P., Fleitout, L. & McKenzie, D. Geoid anomalies and the structure of continental and oceanic lithospheres. *J. Geophys. Res.* **101**, 16119–16135 (1996).
- Pari, G. & Peltier, W. R. The free-air gravity constraint on subcontinental mantle dynamics. *J. Geophys. Res.* **101**, 28105–28132 (1996).
- Peltier, W. R., Forte, A. M., Mitrovica, J. X. & Dziewonski, A. M. Earth's gravitational field: Seismic tomography resolves the enigma of the Laurentian anomaly. *Geophys. Res. Lett.* **19**, 1555–1558 (1992).
- Lenardic, A. & Kaula, W. M. Mantle dynamics and the heat flow into the Earth's continents. *Nature* **378**, 709–711 (1995).
- Bostock, M. G. Ps conversions from the upper mantle transition zone beneath the Canadian landmass. *J. Geophys. Res.* **101**, 8393–8402 (1996).
- Vinnik, L., Chevrot, S. & Montagner, J. P. Shear wave splitting in the mantle Ps phases. *Geophys. Res. Lett.* **24**, 1007–1010 (1997).
- Stammler, K. *et al.* The upper mantle discontinuities: correlated or anticorrelated? *Geophys. Res. Lett.* **19**, 1563–1566 (1992).
- Dueker, K. G. & Sheehan, A. F. Mantle discontinuity structure beneath the Colorado Rocky Mountains and High Plains. *J. Geophys. Res.* **103**, 7153–7169 (1998).
- Chen, Y. H., Roecker, S. W. & Kosarev, G. L. Elevation of the 410 km discontinuity beneath the central Tien Shan: evidence for a detached lithospheric root. *Geophys. Res. Lett.* **24**, 1531–1534 (1997).
- Revenaugh, J. & Jordan, T. H. Mantle layering from ScS reverberations. 2. The transition zone. *J. Geophys. Res.* **96**, 19763–19780 (1991).
- Gossler, J. & Kind, R. Seismic evidence for very deep roots of continents. *Earth Planet. Sci. Lett.* **138**, 1–13 (1996).

20. Kennett, B. L. N. *IASPEI 1991 Seismological Tables* (Research School of Earth Sciences, Australian National University, Canberra, 1991).
21. Mooney, W. D., Laske, G. & Masters, T. G. Crust 5.1: a global crust model at $5^\circ \times 5^\circ$. *J. Geophys. Res.* **103**, 727–747 (1998).
22. Musacchio, G., Mooney, W. D., Luetgert, J. H. & Christensen, N. I. Composition of the crust in the Grenville and Appalachian provinces of North America inferred from Vp/Vs ratios. *J. Geophys. Res.* **102**, 15225–15241 (1997).
23. Isaak, D. G. High-temperature elasticity of iron-bearing olivines. *J. Geophys. Res.* **97**, 1871–1885 (1992).
24. Agnon, A. & Bukowski, M. S. T. δ_s at high pressure and $\ln V_s/\ln V_p$ in the lower mantle. *Geophys. Res. Lett.* **17**, 1149–1152 (1990).
25. Fouch, M. J., Fischer, K. M. & Parmentier, E. M. Mantle flow contributions to shear wave splitting beneath continental roots. *Eos* **78**(17), S322 (1997).
26. Karato, S.-I. & Wu, P. Rheology of the upper mantle: a synthesis. *Science* **260**, 771–778 (1993).
27. Bina, C. R. & Helffrich, G. Phase transition Clapeyron slopes and transition zone seismic discontinuity topography. *J. Geophys. Res.* **99**, 15853–15860 (1994).
28. Chopelas, A. Thermal properties of β - Mg_2SiO_4 at mantle pressures derived from vibrational spectroscopy: implications for the mantle at 400 km depth. *J. Geophys. Res.* **96**, 11817–11829 (1991).
29. Akaogi, M. & Ito, E. Refinement of enthalpy measurement of MgSiO_3 perovskite and negative pressure-temperature slopes for perovskite-forming reactions. *Geophys. Res. Lett.* **20**, 1839–1842 (1993).
30. van der Lee, S. & Nolet, G. Seismic imaging of the subducted trailing fragments of the Farallon plate. *Nature* **386**, 266–269 (1997).

Acknowledgements. We thank S. van der Lee and S. Grand for their mantle velocity models, Y. Shen for assistance with the receiver function analysis, G. Helffrich for a review, the IRIS/PASSCAL program for the seismometers used in the MOMA experiment, and M. Fouch, G. Al-Eqabi, P. Shore and the Lamont PASSCAL Instrument Center for their help with the deployment. Data for stations HRV and CCM were obtained from the IRIS Data Management Center. This work was supported by the Geophysics Program of the US NSF.

Correspondence and requests for materials should be addressed to A.L. (e-mail: aibing@emma.geo.brown.edu).

Allometric scaling of plant energetics and population density

Brian J. Enquist*, James H. Brown* & Geoffrey B. West†

The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

** Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA*

† Theoretical Division, T-8, MS B285, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Scaling relationships that describe variation in population density with body size in ecological communities, such as the thinning law in plant ecology^{1–3}, can be explained in terms of how individuals use resources as a function of their size. Data for rates of xylem transport as a function of stem diameter show that rates of resource use in individual plants scale as approximately the 3/4 power of body mass, which is the same as metabolic rates of animals^{4–7}. Here we use this relationship to develop a mechanistic model for relationships between density and mass in resource-limited plants. It predicts that average plant size should scale as the $-4/3$ power of maximum population density, in agreement with empirical evidence and comparable relationships in animals^{5,6,8}, but significantly less than the $-3/2$ power predicted by geometric models¹. Our model implies that fundamental constraints on metabolic rate are reflected in the scaling of population density and other ecological and evolutionary phenomena, including the finding that resource allocation among species in ecosystems is independent of body size^{5,6,8}.

Many characteristics of organisms vary with body size, as described by allometric equations of the form

$$Y = Y_0 M^b \quad (1)$$

where Y is the dependent variable, M is body mass, b is a power exponent and Y_0 is a normalization constant that varies with the nature of Y and with the kind of organism. Studies of animals suggest that many variables scale with quarter-powers of mass, for example $b \approx 3/4$ for metabolic rate, $-3/4$ for population density and $1/4$ for lifespan^{4–9}. There is now a general model to explain why

many anatomical and physiological scaling exponents of both plants and animals scale as quarter-powers⁷. However, the mechanistic connections between these organismal processes and their ecological consequences remain poorly understood. Allometric studies have traditionally not been an important theme in plant biology⁹. One exception is the relationship between population density and plant size. When the dry mass of the average plant ($\bar{M}$) in mature populations is plotted against the maximum plant density ($N_{\max}$) there is a distinct upper boundary that has traditionally been characterized by a power-law with an exponent of $-3/2$ (refs 1–3). This pattern, known as the thinning law, has been shown to hold for plants in both single- and mixed-species stands and over a size range spanning 23 orders of magnitude from unicellular algae to the tallest trees^{1–3,10}. The fact that a plant fills a volume and covers an area has suggested a simple geometric explanation for the $-3/2$ thinning law¹.

However, the theoretical and empirical bases for the density–mass boundary have been called into question^{11–18}. The $-3/2$ exponent, derived from purely geometric considerations, is difficult to reconcile with known mechanisms of plant growth, resource uptake and competition. Furthermore, increasingly precise data suggest that the boundary is closer to $-4/3$ (refs 14, 17), indicating that population density scales as $M^{-3/4}$, the same as in animals^{5,6,8}. Because the metabolic rates of animals scale as $M^{3/4}$, similar relationships in plants suggest that both share a common scaling law that reflects how resource requirements of individual organisms affect competition and spacing among individuals within ecological communities. Here we apply our general model⁷ for the design of biological resource distribution networks to provide mechanistic connections between resource requirements and population density for plants. The model predicts a fractal-like branching architecture and numerous allometric scaling relationships, for example total leaf mass scales as $M^{3/4}$, trunk diameter as $M^{3/8}$ and resource use, metabolic rate and gross photosynthetic rate as $M^{3/4}$. Although most of these predictions are well supported by data, there have been few allometric studies of total resource use or metabolism of plants.

Our analysis of data in the literature shows that whole-plant resource use scales as $M^{3/4}$. Several studies report the total rate of

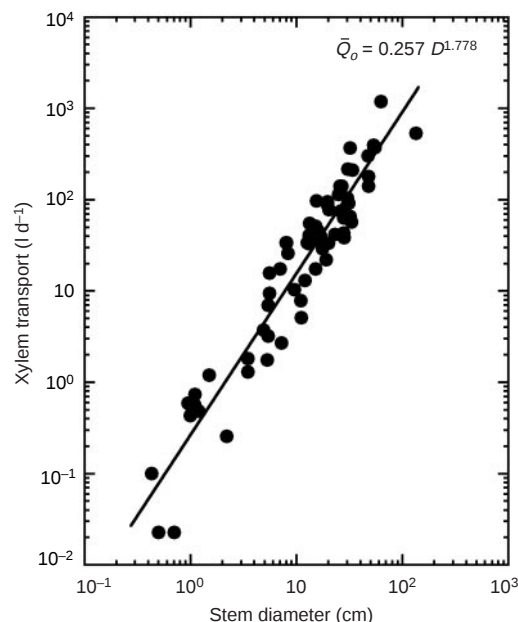


Figure 1 Relationship between the rate of whole-plant xylem transport and basal stem diameter. Data are from 69 individuals and 37 species, including herbaceous plants, shrubs, tree seedlings and mature evergreen and deciduous trees ($r^2 = 0.913$, $n = 69$, $P < 0.0001$; 95% confidence interval: 1.644 to 1.912).

fluid transport in the xylem ($\dot{Q}_0$) as a function of stem diameter (D); this can be described as $\dot{Q}_0 \propto D^{1.778}$ (Fig. 1). Other studies report relationships between stem diameter and above-ground dry mass^{11,19}; averaging these gives $D \propto M^{0.412}$ ($n = 78$, s.d. = 0.356), so that $\dot{Q}_0 \propto M^{0.732}$.

The above relationships are nearly indistinguishable from those predicted from our model: $\dot{Q} \propto D^2$, $D \propto M^{3/8}$ and

$$\dot{Q}_0 \propto M^{3/4}. \quad (2)$$

Small deviations from the predicted exponents can have many sources, including measuring techniques that slightly overestimate the diameter of large trees¹⁵.

Whole-plant xylem transport provides a measure not only of nutrient and water use^{20,21}, but also of gross photosynthesis and therefore of metabolic rate. Because of stoichiometric and physiological constraints, the allometric scaling exponents for water, nutrient and photosynthate fluxes must be equivalent. Thus, rates of transpiration or xylem transport are appropriate, although generally overlooked, indices of plant metabolism. Both the theoretical model and the empirical evidence indicate that whole-plant metabolic rates scale as $M^{3/4}$, so that mass- or tissue-specific rates scale as $M^{-1/4}$. This agrees with the qualitative observation that size-specific growth rates are generally highest in annuals and small herbs and lowest in large trees^{21–24}. More quantitatively, whole-plant rates of twig and leaf production (P_L) and wood and bark production (P_B), in six species of temperate trees, scale as $P_L \propto D^{1.653}$ and $P_B \propto D^{1.807}$ (ref. 25). Using the scaling of diameter in these species ($D \propto M^{0.438}$), the rate of new tissue production is $P_L \propto M^{0.724}$ and $P_B \propto M^{0.791}$, which brackets the predicted value of $M^{3/4}$.

Having established that whole-plant resource uses scales as $M^{3/4}$, we now model the relationship between maximal population density ($N_{\max}$) and average plant mass ($\bar{M}$) in ecological communities. We assume that: (1) sessile plants compete for spatially limited resources; (2) their rate of resource use scales as $M^{3/4}$; and (3) plants grow until they are limited by resources^{21,23,26}. The maximum number of individuals that can be supported per unit area ($N_{\max}$) is related to the rate of resource supply (R) per unit area and the average rate of resource use per individual ($\bar{Q}$) by $R = N_{\max} \bar{Q} \propto N_{\max} \bar{M}^{3/4}$. At equilibrium in any given environment,

R is constant, giving

$$N_{\max} \propto \bar{M}^{-3/4}. \quad (3)$$

This is the form of the allometric equation traditionally used by animal ecologists, who often find a similar $-3/4$ scaling exponent for population density^{5,6,8}. Plant ecologists have traditionally treated mass as the dependent variable, giving $\bar{M} \propto N_{\max}^{-4/3}$.

Our model, based on resource use by individual plants, predicts a mass–density scaling exponent of $-4/3$, rather than $-3/2$ predicted by the geometric model¹. Various studies also suggest that the thinning exponent is close to $-4/3$ (refs 14, 17, 18). Our analysis of data from the literature relating $\bar{M}$ and $N_{\max}$ (Fig. 2) shows that the exponent, -1.341 , has statistical confidence intervals that include $-4/3$ but not $-3/2$. Other sources in the literature express total above-ground plant biomass per unit area (M_{tot}) as a function of maximum population density¹². Using equation (3), our model predicts the scaling of total plant mass

$$M_{\text{tot}} = N_{\max} \bar{M} \propto N_{\max}^{-1/3} \bar{M}^{-1/3}. \quad (4)$$

The geometric model¹ predicts an exponent of $-1/2$. A previous analysis of data on interspecific populations found an exponent of -0.379 (ref. 17), which is closer to $-1/3$ than to $-1/2$. Using a different data set, we performed a similar analysis (Fig. 3) and found that biomass per unit area scales as $N_{\max}^{-0.325}$, which is statistically indistinguishable from the $-1/3$ predicted by our model, but significantly different from the $-1/2$ predicted by the geometric model.

Because the rate of resource use per unit area is the product of population density and the rate of resource use per individual, from equations (2) and (3) we have

$$Q_{\text{tot}} = N_{\max} \bar{Q} \propto \bar{M}^{-3/4} \bar{M}^{3/4} \propto \bar{M}^0. \quad (5)$$

Therefore, total energy use or productivity of plants in ecosystems is predicted to be invariant with respect to body size. We calculated Q_{tot} from the data used to compile Figs 1 and 2. As shown in Fig. 4, the rate of resource use per surface area scales as $\bar{M}^{0.014}$. This empirical value does not differ statistically from the size invariance ($\bar{M}^0$) predicted by the model. The relationship holds across variation of 12 orders of magnitude in plant size. The variation around

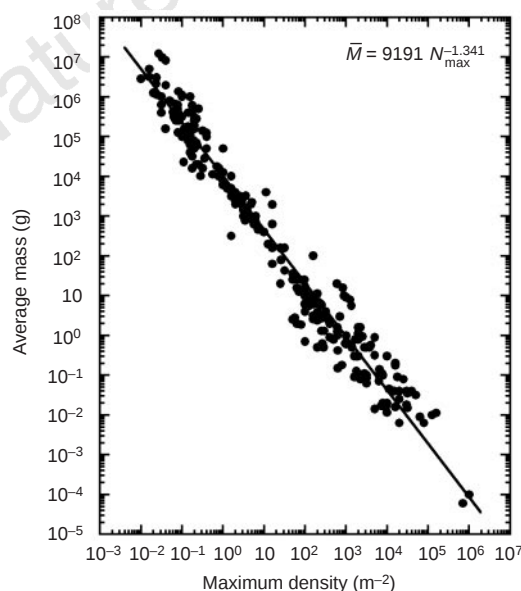


Figure 2 Relationship between average plant mass and maximum population density. Plant sizes range from *Sequoia* to *Lemna*. Data are from 251 populations ($r^2 = 0.963$, $n = 251$, $P < 0.0001$; 95% confidence interval: -1.374 to -1.309). The exponent is statistically indistinguishable from $-4/3$, indicating that maximum population density scales as $M^{-3/4}$.

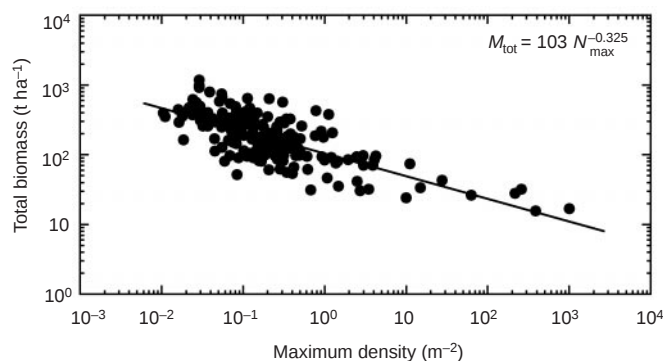


Figure 3 Relationship between total plant biomass (roots, shoots and leaves) and maximal population density²⁰. ($r^2 = 0.9534$, $n = 189$, $P < 0.0001$; 95% confidence interval: -0.281 to -0.368). The fitted equation has an exponent statistically indistinguishable from $N_{\max}^{-1/3}$ predicted from our model.

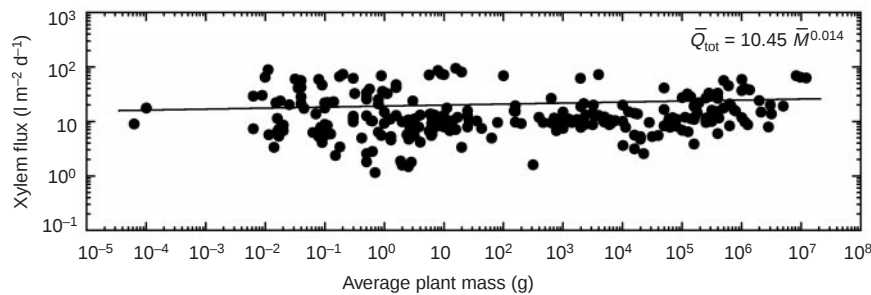


Figure 4 Relationship between total xylem flux and average size of the dominant plants in diverse ecosystems. ($r^2 = 0.009$, $n = 251$, $P < 0.127$; 95% confidence interval: -0.004 to 0.031). The stoichiometric equivalence of rates of water, nutrient and photosynthate flux within individual plants (see text) explains why evapo-

transpiration can be used to estimate productivity²⁷. Because the allometric equation has an exponent that is statistically indistinguishable from zero, ecosystem productivity is independent of plant size.

the regression line reflects variation in resource supply, and therefore in productivity among ecosystems ranging from arid grasslands and tundra to temperate and tropical forests²⁷. Most of the dissatisfaction with the original geometric formulation of the thinning law stems from its lack of a means of predicting empirically measured thinning trajectories, and the resulting variation in sizes and densities of plants in different environments^{3,11–18}. For example, as plant populations thin, individuals attain a certain maximum size that is characteristic of local environmental conditions. Not only are individuals larger in forests than in grasslands, but also each community typically contains multiple codominant species of nearly identical size²³. Our model does not predict thinning trajectories, but it does predict that the rate of resource use per unit area varies among plant communities with differences in resource supply but not with plant size. Thus, ecosystems composed of plants of contrasting sizes and life forms, such as certain forests, grasslands and agricultural fields, can have identical productivity²⁷.

We have shown that metabolic rates of plants scale indistinguishably from the predicted $M^{3/4}$, and that a model incorporating this scaling can account for the maximum sizes and densities of plants observed in different communities. Traditionally, plant ecologists have implicitly treated the sizes of individuals as if they were determined by population density, by plotting mass as the dependent variable in depicting thinning relationships. Animal physiologists and ecologists have done just the opposite, plotting density and other variables as functions of body mass. The theoretical and empirical advantage of the latter approach is that variables are expressed in terms of standard allometric equations, such as equation (1), which highlight the universality of the 3/4-power scaling of resource use and the related 1/4-power scaling of other structural, functional and ecological attributes^{4–8}. It has long been known that life-history variables, such as growth rate, lifespan and age to first reproduction, scale with body size^{21–24,28}. Despite this variation, rates of resource use per unit area are independent of body size. This is observed empirically in animals⁸ and demonstrated here, both empirically and theoretically, for vascular plants. A common body of allometric theory promises to provide a general framework for explaining many features of biological diversity. □

Methods

Data from the literature were compiled, plotted and analysed to define the allometric relationships shown in Figs 1–4. Data and their sources are available from B.J.E. The allometric equations were fitted to log-transformed data by least-squared regression. For simplicity, scaling relationships are presented in the text with normalization constants omitted. Data of maximum whole-plant xylem transport (Fig. 1) came from measurements using heat balance²⁰ or radioactive tracers²⁹. Data for maximum values of average plant mass and plant density (Fig. 2) were obtained from refs 1, 2, 13, 14, 17 and additional studies. Data for total plant biomass (Fig. 3) were taken from ref. 30. Values were

obtained from diverse ecosystems with woody vegetation from around the world but excluded forests with heavy epiphyte loads. Values for total ecosystem xylem flux (Fig. 4) were calculated by taking the values of plant population density and average plant mass used in Fig. 2, and multiplying by the estimated rate of xylem transport using the equation in Fig. 1. Normalization constants were incorporated¹⁹ to express values as rates of material transported from the soil and moved vertically through plant stems (in units of $l m^{-2} d^{-1}$).

Received 13 March; accepted 7 July 1998

- Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. Self-thinning in overcrowded pure stands under cultivated and natural conditions. *J. Biol. Osaka City Univ.* **14**, 107–129 (1963).
- Gorham, E. Shoot height, weight and standing crop in relation to density of monospecific stands. *Nature* **279**, 148–150 (1979).
- White, J. in *Studies on Plant Demography: A Festschrift for John L. Harper* (ed. White, J.) 291–301 (Academic, New York, 1985).
- Schmidt-Nielsen, K. *Scaling: Why is Animal Size so Important?* (Cambridge Univ. Press, 1984).
- McMahon, T. A. & Bonner, J. T. *On Size and Life* (Scientific American Library, New York, 1983).
- Brown, J. H. *Macroecology* (Univ. Chicago Press, 1995).
- West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126 (1997).
- Damuth, J. Population density and body size in mammals. *Nature* **290**, 699–700 (1981).
- Niklas, K. J. *Plant Allometry: The Scaling of Form and Process* (Univ. Chicago Press, 1994).
- Agusti, S., Duarte, C. M. & Kalf, J. Algal cell size and the maximum density and biomass of phytoplankton. *Limnol. Oceanogr.* **32**, 983–986 (1987).
- White, J. The allometric interpretation of the self-thinning rule. *J. Theor. Biol.* **89**, 475–500 (1981).
- Petratis, P. S. Use of average vs. total biomass in self-thinning relationships. *Ecology* **76**, 656–658 (1995).
- Weller, D. E. A reevaluation of the $-3/2$ power rule of plant self-thinning. *Ecol. Monogr.* **57**, 23–43 (1987).
- Weller, D. E. The interspecific size-density relationship among crowded plant stands and its implications for the $-3/2$ power rule of self-thinning. *Am. Nat.* **133**, 20–41 (1989).
- Norberg, R. A. Theory of growth geometry of plants and self-thinning of plant populations: geometric similarity, elastic similarity, and different growth modes of plant parts. *Am. Nat.* **131**, 220–256 (1988).
- Osawa, A. & Allen, R. B. Allometric theory explains self-thinning relationships of mountain beech and red pine. *Ecology* **74**, 1020–1032 (1993).
- Lonsdale, W. M. The self-thinning rule: dead or alive? *Ecology* **71**, 1373–1388 (1990).
- Franco, M. & Kelly, C. K. The interspecific mass-density relationship and plant geometry. *Proc. Natl Acad. Sci. USA* **95**, 7830–7835 (1998).
- Smith, W. B. & Brand, G. J. *Allometric Biomass Equations for 98 Species of Herbs, Shrubs, and Small Trees* (North Central Forest Experimental Station Research Note NC-299, Forest Service—USDA, 1983).
- Kozlowski, T. T. & Pallardy, S. G. *Physiology of Woody Plants* (Academic, New York, 1997).
- Chapin, S. F. III in *Scaling Physiological Processes: Leaf to Globe* (ed. Ehleringer, J. R.) 287–319 (Academic, New York, 1993).
- Grime, J. P. & Hunt, R. Relative growth-rate: its range and adaptive significance in a local flora. *J. Ecol.* **63**, 393–422 (1975).
- Tilman, D. *Plant Strategies and the Dynamics and Structure of Plant Communities* (Princeton Univ. Press, 1988).
- Yoda, K., Shinozaki, K., Ogawa, J., Hozumi, K. & Kira, T. Estimation of the total amount of respiration in woody organs of trees and forest communities. *J. Biol. Osaka City Univ.* **16**, 15–26 (1965).
- Whittaker, R. H. & Woodwell, G. M. Dimension and production relations of trees and shrubs in the Brookhaven Forest, New York. *Ecology* **56**, 1–25 (1968).
- Huston, M. A. & DeAngelis, D. Competition and coexistence: the effects of resource transport and supply rates. *Am. Nat.* **144**, 954–977 (1994).
- Rosenzweig, M. L. Net primary productivity of terrestrial communities: prediction from climatological data. *Am. Nat.* **102**, 67–74 (1968).
- Charnov, E. L. *Life History Invariants: Some Explorations of Symmetry in Evolutionary Ecology* (Oxford Univ. Press, 1993).
- Shulze, E. D. et al. Canopy transpiration and water fluxes in the xylem of the trunk of *Larix* and *Picea* trees—comparison of xylem flow, porometer and cuvette measurements. *Oecologia* **66**, 475–483 (1985).
- Cannell, M. G. R. *World Forest Biomass and Primary Production Data* (Academic, New York, 1982).

Acknowledgements. We thank G. C. Stevens, C. A. F. Enquist, H. S. Horn, T. K. Lowrey, D. Marshall, K. J. Niklas, J. T. Bonner and J. Damuth for their help. B.J.E. was supported by a Fulbright fellowship and funding by NSF; J.H.B. by NSF; and G.B.W. by the US Department of Energy.

Correspondence and requests for materials should be addressed to B.J.E. (e-mail: benquist@unm.edu).

Prism adaptation to a rightward optical deviation rehabilitates left hemispatial neglect

Yves Rossetti^{†*}, Gilles Rode^{†*}, Laure Pisella[†],
Alessandro Farnè[†], Ling Li^{*}, Dominique Boisson^{*‡}
& Marie-Thérèse Perenin[†]

^{*} Service de Rééducation Neurologique, Hôpital Henry Gabrielle, Hospices Civils de Lyon and Université Claude Bernard, BP 57, 69565 St Genis Laval Cedex, France

[†] INSERM, Unité 94: Espace et Action, 16 avenue Lépine, Case 13, 69676 Bron, France

[‡] Service de Rééducation, Hôpital Neurologique, 59 Bd Pinel, 69003 Lyon, France

A large proportion of right-hemisphere stroke patients show hemispatial neglect—a neurological deficit of perception, attention, representation, and/or performing actions within their left-sided space¹, inducing many functional debilitating effects on everyday life, and responsible for poor functional recovery and ability to benefit from treatment². The frequent parietal locus of the lesion producing neglect reflects the impairment of coordinate transformation used by the nervous system to represent extra-personal space. Given that adaptation to a visual distortion can provide an efficient way to stimulate neural structures responsible for the transformation of sensorimotor coordinates, the aim of our study was to investigate the effect of prism adaptation on various neglect symptoms, including the pathological shift of the subjective midline to the right. All patients exposed to the optical shift of the visual field to the right were improved on their manual body-midline demonstration and on classical neuropsychological tests. Unlike other physiological manipulations used to improve neglect, this improvement lasted for at least two hours after prism removal and thus could be useful in rehabilitation programmes. The positive effect found for both sensorimotor and more cognitive spatial functions suggests that they share or depend on a common level of space representation linked to multisensory integration.

The manifestations of hemispatial neglect may be temporarily improved by several sensory manipulations (caloric stimulation^{3,4}, neck vibration⁵ or optokinetic stimulation⁶), whereas other manipulations have failed to produce a consistent improvement of neglect⁷. Despite over twenty years of extensive research efforts, however, in all such cases the effects are usually transitory, lasting no more than 10–12 minutes. One classical feature of neglect is a pathological shift of the subjective midline to the right, as measured by straight-ahead demonstration in the dark⁸. Psychophysical manipulations can be used to alter straight-ahead demonstrations in normal healthy subjects. Exposure to an optical alteration of the visual field is known to produce an initial disorganization of visuo-motor behaviour, which can be corrected through visuo-motor adaptation⁹. Such adaptation has been widely used to demonstrate the plasticity of coordinate transformations involved in multi-sensory and sensori-motor integration^{10,11}. One major compensative effect of short-term wedge-prism exposure is a shift of proprioceptive representations, which can be demonstrated by asking subjects to point straight ahead in the dark (see ref. 9, for example). After adaptation, this finger straight-ahead demonstration is shifted in a direction opposite to the optical deviation, indicating that internal visual and proprioceptive 'maps' have been realigned. The aim of our study was thus to investigate the effect of prism adaptation on various neglect symptoms, including the manual demonstration of the subjective midline.

Sixteen right brain-damaged patients with a persistent, left hemispatial neglect participated in the study (nine males and seven females; mean age, 62). All had been admitted to a neurolo-

gical rehabilitation hospital for a moderate to severe hemiplegia and various degrees of somatosensory dysfunction (see Table 1 of Supplementary Information). All patients were right-handed and had a documented, single unilateral hemispheric lesion, and no past history of previous stroke. None of the patients suffered from impaired vigilance, confusion, general mental deterioration or psychiatric disorders. Testing took place between 3 weeks and 14 months post-onset (average, 9 weeks).

Experiment 1 was aimed at measuring the adaptability of neglect patients to a lateral shift of the visual field. Manual body-midline demonstration was used to evaluate adaptation to base-left wedge prisms (inducing a 10° shift of the visual field to the right) by a simple target-pointing task. A group of eight neglect patients and a group of five control subjects produced 10 straight-ahead pointing trials before and after a short period of adaptation. Figure 1 shows that the patient's mean straight-ahead was initially shifted to the right. Following the adaptation, both patients and controls demonstrated straight-ahead shifts to the left, and thus the patient's pathological deviation was greatly improved. This first result demonstrated that neglect patients can easily adapt to a lateral shift of the visual field to the right, and that prism adaptation, acting against the rightward bias of straight-ahead demonstration (on average), allows these patients to show a close-to-normal post-test performance. (Interestingly, patients did not adapt to a shift of the visual field to the left, whereas normals adapted with the same amount to the left and the right shift.)

Experiment 2 investigated whether prism adaptation could also improve the main clinical manifestations of neglect¹². Twelve neglect patients were randomly assigned to the prism group and the control group. All patients underwent a similar procedure. A series of traditional neuropsychological tests² was performed on three sessions. After the first session (pre-test), patients had to perform the elementary pointing task with the prismatic goggles used in experiment 1, in order to adapt to the visual-proprioceptive discrepancy. Patients in the 'prism' group were exposed to the 10-degree optical deviation to the right. Patients in the control group

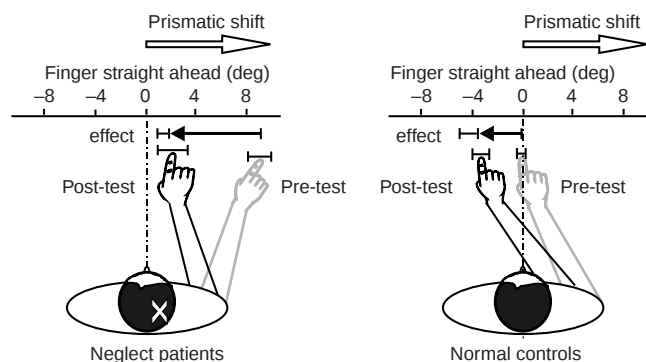


Figure 1 Midline demonstrations. Blindfolded subjects were required to point straight ahead while their head was kept aligned with the body's sagittal axis. Ten pointing trials were run in the pre-test (without goggles) and in the post-test (immediately upon removal of the 10-deg prism (white arrow)). As expected, the midline demonstrations made by the neglect group were initially shifted to the right, whereas control subjects pointed to their actual straight ahead. A two-way repeated-measures analysis of variance with the factors Group ('neglect' vs controls) and Session (pre-test vs post-test) was performed on midline demonstrations (mean of ten responses) obtained for each subject in each condition, and showed a main effect of Session ($P < 0.05$) and a Group by Session interaction ($P < 0.05$) indicating that the two groups were not affected by the prisms in the same way. Additional comparison showed that the differential effect (pre-test minus post-test) was significantly higher in the neglect group ($F(1, 11) = 8.70$; $P < 0.05$). Patients were thus more affected by the adaptation than controls (black arrows), and the magnitude of this effect was less variable in the patients (arrow extensions).

wore neutral goggles with thick flat glasses. Immediately after removing the goggles, a second session was performed with the same battery of tests (post-test). Patients were again tested after a delay of about two hours following the goggles exposure (late test). The standard neuropsychological procedure included: line bisection¹³, line cancellation¹⁴, copying a simple drawing made of five items¹⁵, drawing of a daisy from memory, and reading a simple text. All patients in the 'prism' group exhibited a clear improvement following prism exposure. A Group by Session interaction effect ($F(2,8)$, $P < 0.01$) showed that 'prism' and control patients were differentially affected by the goggles exposure. Post-hoc Sheffé's tests demonstrated that the post-test and the late test provided significantly better scores than the pre-test (both $P < 0.05$) in the 'prism' group, whereas no significant difference was found between testing sessions for the control group (both $P > 0.90$). Figure 2 shows results obtained for the Gainotti copying test¹⁵ by one representative patient of each group and the two group values. Figure 3 allows further group comparison for cancellation and bisection tasks. A dramatic improvement was seen in all tests following prism exposure and was fully maintained two hours later. By contrast, there was no significant improvement in the control group.

Unlike other physiological manipulations used to improve neglect, patients demonstrating the classical features of severe visual neglect after right-brain damage showed a strong and reliable

improvement after a short adaptation period to a prismatic shift of the visual field to the right. These results indicate that the process of prism adaptation, long considered as a relatively passive process involved in the recalibration of visuo-motor coordination, can also affect the organization of higher levels of spatial representation, such as those impaired in neglect patients. This finding also stands in contrast with the usual specificity of wedge prisms after-effects in normal subjects, found to be arm-specific or even task-specific under several experimental conditions. It should stimulate the search for higher-level after-effects in healthy subjects as well as investigations about why a patient with a lesion of the parietal network may be more permeable to prism adaptation than normals. If prism adaptation can influence higher-level spatial representations, then this result raises an intriguing question regarding the putative causal role of subjective midline shift on the other common features of neglect. In the current debate, such an explanation has been put forward (see ref. 5, for example) but has been challenged by experiments providing evidence for the lack of a consistent relationship between neglect and a change in the egocentric reference^{16,17}. The co-occurrence of a positive effect of prism adaptation on the manual straight-ahead (independent of vision) and on conventional tests of visual neglect suggests that a common level of space representation (linked to multisensory integration) may be shared in both sensorimotor (for example, straight-ahead pointing) and more cognitive spatial functions (such as copying, reading). Whether prism adaptation can also improve such higher functions as mental imagery or anosognosia remains to be investigated.

This robust effect of prism adaptation cannot be attributed solely to an effect of attention, nonspecific activation of the right lesioned hemisphere, or improvement of defective left-sided sensory processes, all of which are factors that have been proposed to account for the effect of sensory stimulation on neglect¹⁸. Rather, it would seem that the improvement of patients depended upon a central mechanism, which may be mediated by a left hemisphere activation¹⁹. Therefore the effect of the prism is not merely to act as a passive, prosthetic modulator, modifying sensory afferents (see refs 18, 20 for example), but can be conceived of as stimulating active processes involved in the plasticity of sensorimotor correspondences, by activating brain functions related to multisensory integration and space representation. In this respect, our results contrast with those obtained when Fresnel prisms were used to shift the left visual field towards the central retina, rather than to study the effect of adaptation²⁰. Under these latter conditions, no significant improvement of neuropsychological testing was observed after the patients kept using the goggles daily during four weeks. A permanent exposure to a shift of the visual field in a direction compensating for the left hemispatial neglect is thus less effective on neglect than the adaptation to this optical deviation, because it produces mainly a sensory, peripheral effect rather than a central effect on space representation.

Two non-exclusive mechanisms can be proposed to account for the strong improvement of neglect after prism adaptation: a general plasticity induction and a lateralized warning signal. First, stimulating the short-term plasticity of brain functions related to coordinate transformations and space representation may favour the neural restoration of the right-hemisphere functions when they have been impaired by a lesion. Indeed, adaptation to visual field reversal is able to induce functional reorganization of early visual processing¹¹.

Second, a direction-specific mechanism has to be evoked to account for the direction-specific alteration of the egocentric reference seen in experiment 1, and for the overcompensation of neglect observed with the bisection task in experiment 2 (Fig. 3; Schenkenberg test, right side). There is a natural tendency of the brain to compensate for distortions occurring at either the sensory⁹ or the motor side²¹. The natural velocity of such processes strongly contrasts with the very slow spontaneous recovery classically observed for hemispatial neglect. Whereas neglect patients are

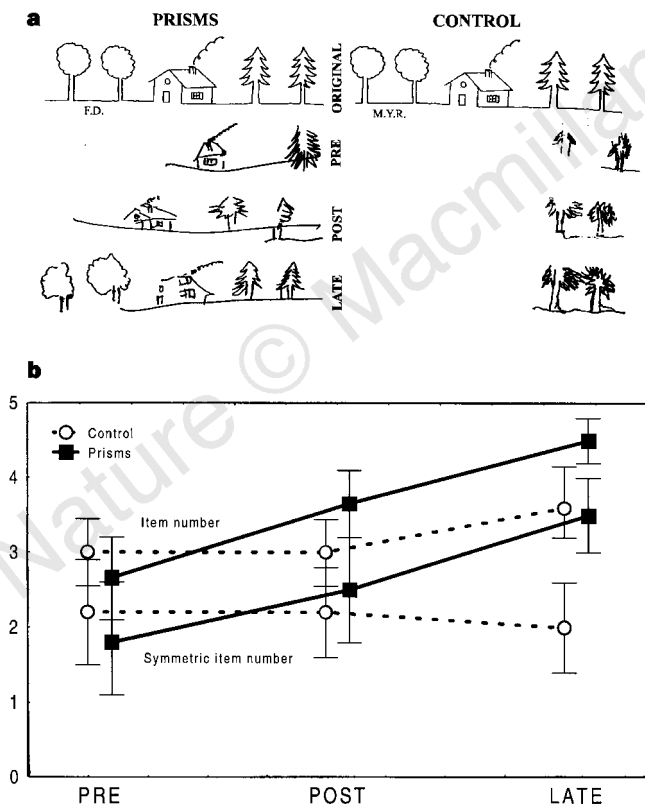


Figure 2 Copying test. One of the classic tests used to assess hemispatial neglect is referred to as the Gainotti test¹⁵, in which the patient is required to copy a drawing made of five items. **a**, A representative example from patient F.D. (left panel) in which the drawing made before prism exposure (pre-test) demonstrates the complete neglect of three items. In the post-test (on prism removal), one item is added to the patient's drawing; in the late test (after 2 h), all items are drawn. By contrast, patient M.Y.R. (right panel) was exposed to neutral goggles and she did not improve in the post and late tests. **b**, The mean number of items drawn (± s.e.m.) and the mean number of items drawn symmetrically (± s.e.m.) in the two groups. Note that these two scores, respectively reflecting space-based and object-based hemineglect, are improved in the same way in the 'prisms' group. No decrease in performance was seen between the post-test and the late test in this particular test, or in others.

typically biased towards the right side, they usually seem to maintain some coherence between various aspects of brain functions related to space, which could prevent them from compensating their biases. The coherent disorganization of space representation in neglect does not seem to produce effective disparity signals, whereas the visual-proprioceptive discrepancy induced by prism exposure alters the coordinate transformations used by the nervous system to represent extrapersonal space. The dramatic improvement induced by prism adaptation suggests that a signal is given to the brain that stimulates the natural recovery process. The main signal for this adaptation is the discrepancy observed between the expected position of the hand and its shifted, seen position as it enters in the visual field. This signal precisely indicates to the subject that his actual action is biased towards the right as compared with his intention. The lateralized information introduced by prism exposure may break the biased coherence and introduce a signal useful for stimulating the correction of left neglect. This should also explain why the neglect patients showed a stronger adaptation than the controls in experiment 1, and a prolonged effect. (Preliminary results suggest that the effect of prism adaptation on neglect may last over four days for the two types of tasks used here.)

Neural structures considered to be involved in prism adaptation have long been restricted to the cerebellar region (reviewed in ref. 22). However, the posterior parietal cortex contralateral to the acting hand might be activated during adaptation to a prism-induced shift of the visual field¹⁹. Therefore the question arises of the respective contributions of these two brain structures in the sensory realignment and in the more strategic aspects of prism adaptation. Intraparietal areas are well known for being involved in

sensorimotor transformations²³. These early transformations could provide the basis for higher-level spatial representations, including the level of sensorimotor interface required by action planning²⁴.

An attractive aspect of prism exposure lies in its non-invasive nature, its acceptability to patients, and its ease of use. In addition, the duration of the effects described here, owing to central active processes stimulated by the adaptation, indicates that this technique may come to have a major role in the neuropsychological rehabilitation of hemispatial neglect. □

Methods

Prism adaptation. The prism exposure condition was identical in the two experiments. The goggles were fitted with wide-field, prismatic lenses, creating an optical shift of 10 deg. The exposure period consisted in making 50 pointing responses to visual targets presented 10 deg to the right or left of the objective body midline. During the prism exposure, subjects were asked to point at a fast but comfortable speed; they could see the target, the second half of their pointing trajectory and their terminal error. Their head was kept aligned with the body's sagittal axis by a chin-rest and controlled by an investigator. The duration of this exposure ranged from 2 to 5 min. By contrast with a previous study, which used a shift of only the left half of the visual field to draw attention to the neglected field¹³, our experiment evaluated the effect of prism adaptation, and all subjects were subsequently tested without prisms.

Subjects. All patients and subjects were right-handed. Patients presented a unique lesion of the right hemisphere, sparing the cerebellum and other brainstem structures. On computed tomography (CT) scan, 11 cases presented with a cortico-subcortical hypodense lesion involving the right parieto-temporo-occipital carrefour; three cases with a subcortical lesion that included the right capsulo-lenticular region, and two cases with a right frontal lesion sparing the parietal lobe. Neurological examination revealed no optic ataxia, cerebellar dysfunction, somatoparaphrenia. Five patients initially showed an anosognosia for hemiplegia and hemianopia, which had disappeared at the time of testing. Nine patients presented a left hemianopia on the Goldman testing. Patients participating to both experiment 1 and experiment 2 were enrolled with a minimal time interval of one week and in a random order.

Control subjects participating in experiment 1 were inpatients with no neurological history, matched for age (mean: 59.6 years), sex (3/2) and handedness.

Experiment 1. All subjects participating in experiment 1 were randomly subjected to two different wedge prism exposure conditions (base left, base right) with a minimal time interval of two days. The order of presentation was counterbalanced. They were seated blindfolded in front of a horizontal box that permitted measurement of the finger movement endpoints with an accuracy of 1 deg. The cover of the experimental box allowed them to perform free pointing movements without vision of the arm, or with a partial sight of the arm trajectory (exposure condition).

Experiment 2. Five classical tests were used in experiment 2. All testing was carried out in the midsagittal plane. The line-bisection test allowed the computation of three scores, reflecting the bias in bisecting horizontal lines in the right, centre and left part of the testing sheet. The line-cancellation test provided two scores that reflected omissions made in the right and left halves of the page. Two scores were extracted from the copying test, respectively reflecting the number of items drawn by the patient and the number of items drawn symmetrically. Drawings of a daisy made by each patient were ranked from 1 to 3 in order of symmetry quality by two independent judges. The number of word omissions or alterations was counted for each reading performed. This series of tests thus provided nine scores for each session and for each patient. None of these nine scores showed a significant difference between the two groups of patients during the pre-test (all $P \geq 0.20$).

Received 14 April; accepted 17 June 1998.

1. Heilman, K. M., Watson, R. T. & Valenstein, E. in *Clinical Neuropsychology* (eds Heilman, K. M. & Valenstein, E.) 243–293 (Oxford University Press, New York, 1985).
2. Halligan, P. W., Marshall, J. C. & Wade, D. T. Visuospatial neglect: underlying factors and test sensitivity. *Lancet* **2**, 908–911 (1989).
3. Rubens, A. B. Caloric stimulation and unilateral visual neglect. *Neurology* **35**, 1019–1024 (1985).
4. Rode, G. & Perenin, M. T. Temporary remission of representational hemineglect through vestibular stimulation. *Neuroreport* **5**, 869–872 (1994).
5. Karnath, H. O. Subjective body orientation in neglect and the interactive contribution of neck muscle proprioception and vestibular stimulation. *Brain* **117**, 1001–1012 (1994).

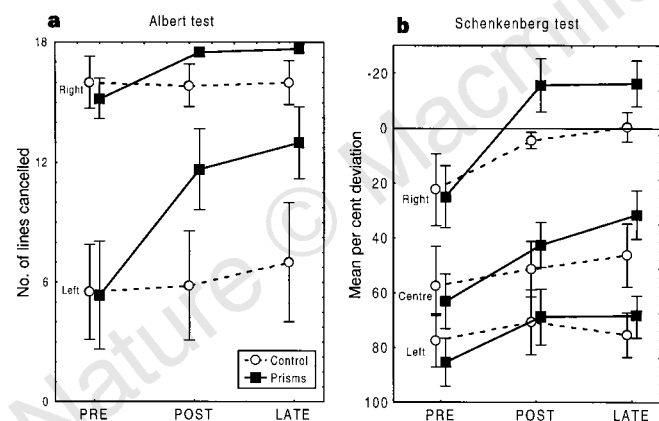


Figure 3 Line cancellation and bisection. **a**, Results obtained by the two groups on a line cancellation task¹⁴ for the two sides of the test sheet (means $\pm$ s.e.m.). The results were classic in that patients performed initially less well in the left half than in the right half (pre-test). The 'prism' group is dramatically improved compared to the control group; this positive effect is not reduced after two hours. The analysis of variance performed for this particular test showed a Group by Session interaction ($F(2,20) = 5.98$; $P < 0.05$), and the post-hoc Sheffé's tests revealed that the pre-test differed from the post-test and the late test (both $P < 0.01$), whereas this was not the case in the control group ($P < 0.95$ for both). A difference between the post-test and the late test was not found in either group (both $P > 0.95$). **b**, Results of the two groups in a line bisection task¹³ for the three parts of the test sheet. The mean deviation ($\% \pm$ s.e.m.) is initially shifted to the right in both groups: patients tend to omit lines on the left side (values close to 100%). Following exposure while wearing goggles, this initial deviation is shifted to the left in the 'prism' group. The analysis of variance showed a Group by Session interaction ($F(2,20) = 5.23$; $P < 0.05$), and the post-hoc Sheffé's tests revealed that the pre-test differed from the post-test and the late test (both $P < 0.01$), whereas this was not the case in the control group (for both, $P > 0.25$). Again, no group exhibited a difference between the post-test and the late test (for both, $P > 0.95$). Note that the post and late values obtained in the right side show a deviation to the left, which strongly supports a directional effect of prism adaptation on space representation in neglect patients.

6. Pizzamiglio, L., Frasca, R., Guariglia, C., Incoccia, C. & Antonucci, G. Effects of optokinetic stimulation in patients with visual neglect. *Cortex* **26**, 535–540 (1990).
7. Walker, R., Young, A. W. & Lincoln, N. B. Eye patching and the rehabilitation of visual neglect. *Neuropsychol. Rehab.* **6**, 219–231 (1996).
8. Perenin, M. T. in *Parietal Lobe Contribution to Orientation in 3D Space* (eds Thier, P. & Karnath, H. O.) 298–308 (Springer, Heidelberg, 1997).
9. Redding, G. M. & Wallace, B. Adaptive spatial alignment and strategic perceptual-motor control. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 379–394 (1996).
10. Rossetti, Y., Koga, K. & Mano, T. Prismatic displacement of vision induces transient changes in the timing of eye-hand coordination. *Percept. Psychophys.* **54**, 355–364 (1993).
11. Sugita, Y. Global plasticity in adult visual cortex following reversal of visual input. *Nature* **380**, 523–526 (1996).
12. Halligan, P. W. & Cockburn, J. M. Cognitive sequelae of stroke: visuospatial and memory disorders. *Crit. Rev. Phys. Rehab. Med.* **5**, 57–81 (1993).
13. Schenkenberg, T., Bradford, D. C. & Ajax, E. T. Line bisection with neurologic impairment. *Neurology* **30**, 509–517 (1980).
14. Albert, M. L. A simple test for visual neglect. *Neurology* **23**, 658–664 (1973).
15. Gainotti, G., Messeri, P. & Tisot, R. Qualitative analysis of unilateral spatial neglect in relation to laterality of cerebral lesions. *J. Neurol. Neurosurg. Psychiatr.* **35**, 545–550 (1972).
16. Chokron, S. & Bartolomeo, P. Patterns of dissociation between left hemineglect and deviation of the egocentric reference. *Neuropsychologia* **35**, 1503–1508 (1997).
17. Farné, A., Ponti, F. & Ladavas, E. In search of biased egocentric reference frames in neglect. *Neuropsychologia* **247**, 611–623 (1998).
18. Vallar, G., Guariglia, C. & Rusconi, M. L. in *Parietal Lobe Contribution to Orientation in 3D Space* (eds Thier, P. & Karnath, H. O.) 555–579 (Springer, Heidelberg, 1997).
19. Clower, D. M. et al. Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature* **383**, 618–621 (1996).
20. Rossi, P. W., Kheifets, S. & Reding, M. J. Fresnel prisms improve visual perception in stroke patients with homonymous hemianopia or unilateral visual neglect. *Neurology* **40**, 1597–1599 (1990).
21. Coello, Y., Orliaguet, J.-P. & Prablanc, C. Pointing movement in an artificial perturbing inertial field: a prospective paradigm for motor control study. *Neuropsychologia* **34**, 879–892 (1996).
22. Jeannerod, M. & Rossetti, Y. in *Visual Perceptual Defects* (ed. Kennard, C.) 439–460 (Tindall, London, 1993).
23. Sakata, H. & Taira, M. Parietal control of hand action. *Curr. Opin. Neurobiol.* **4**, 847–856 (1994).
24. Mattingley, J., Hussain, M., Rorden, C., Kennard, C. & Driver, J. Motor role of human inferior parietal lobe revealed in unilateral neglect patients. *Nature* **392**, 179–182 (1998).

Supplementary information is on www.nature.com. Paper copies are available from the London editorial office of *Nature*.

Acknowledgements. We thank M. Cabanac, D. Clower, L. Goffart, P. Halligan, J. Mattingley, C. Prablanc, A. Vighetto for their comments on the manuscript, and P. Monjaud and C. Urquizar for technical assistance. This work was supported by the Région Rhône-Alpes, France, by the Hospices Civils de Lyon and by Elyses. A.F. was awarded a predoctoral fellowship by the European Neuroscience Programme, European Science Foundation.

Correspondence and requests for materials should be addressed to Y.R. (e-mail: rossetti@lyon151.inserm.fr).

Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness

Ian H. Robertson*, Jason B. Mattingley*, Chris Rorden*† & Jon Driver†

* MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK

† Institute of Cognitive Neuroscience, University College London, London WC1E 6BT, UK

Patients with extensive damage to the right hemisphere of their brain often exhibit unilateral neglect of the left side of space. The spatial attention of these patients is strongly biased towards the right¹, so their awareness of visual events on the left is impaired². Extensive right-hemisphere lesions also impair tonic alertness (the ability to maintain arousal)^{3–5}. This nonspatial deficit in alertness is often considered to be a different problem from spatial neglect^{5,6}, but the two impairments may be linked^{7,8}. If so, then phasically increasing the patients' alertness should temporarily ameliorate their spatial bias in awareness. Here we provide evidence to support this theory. Right-hemisphere-neglect patients judged whether a visual event on the left preceded or followed a comparable event on the right. They became aware of left events half a second later than right events on average. This spatial imbalance in the time course of visual awareness was corrected when a warning sound alerted the patients phasically. Even a warning sound on the right accelerated the perception of left visual events in this way.

Nonspatial phasic alerting can thus overcome disabling spatial biases in perceptual awareness after brain injury.

Attention has at least two components: nonspatial alertness and spatial selectivity^{3,5,9,10}. Alertness is a state of general readiness^{11–13}, whereas selective attention operates to enhance perception at particular spatial locations¹⁴. Spatial attention may depend on a predominantly right-lateralized cortical network^{15,16} that becomes unbalanced following large lesions centred on the right parietal lobe. The resulting stronger activation of the left hemisphere rather than the right hemisphere in the lesioned network biases spatial attention rightwards, producing left neglect^{17,18}. Functional imaging data¹⁹ show that sustaining alertness also relies on a right-lateralized cortical network, involving frontal and parietal lobes. Consistent with this, extensive right-hemisphere damage produces a nonspatial deficit in maintaining alertness^{4,5,7,8,20–22}. Although the effortful maintenance of tonic alertness over prolonged periods relies on this cortical, right frontal–parietal circuit (which is damaged in neglect patients with extensive right-hemisphere lesions), phasic alerting by salient external events depends instead on ascending thalamic–mesencephalic projections (which should remain intact in such patients)^{7,8,10,16,19,23,24}. Hence, it should be possible to phasically activate what remains of the right-lateralized cortical networks for attention and alertness¹⁰ in neglect patients, through alerting events (for example, warning tones). This should effectively shift spatial attention leftwards in neglect patients^{7,8}, thus compensating for their deficit.

Here we tested this theory by playing occasional warning sounds to phasically alert left-neglect patients, as they performed a task that directly measures their speed of awareness for left visual events relative to right events^{2,25}. We predicted that the warning sound should accelerate their perception of left visual events, by shifting attention leftwards because of phasic activation of the damaged right-hemisphere through the intact thalamic–mesencephalic ascending circuit. In every trial, a central fixation cross was presented, followed by a visual bar on either the left or right (see Methods, Fig. 1). This visual bar was followed, at a variable stimulus onset asynchrony (SOA), by a second bar on the other side. The patient had to judge which bar came first. The patients were 'warned' in a random 25% of trials, by a 300-ms central-tone

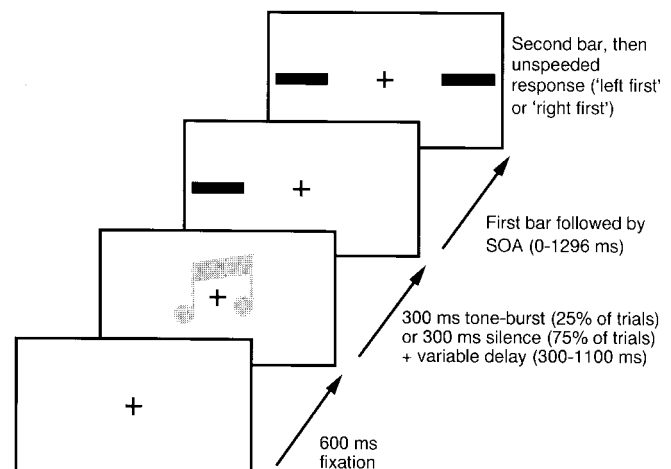


Figure 1 The sequence of events for a single trial, with time running from bottom left to top right. The trial began with the patient fixating the central cross for 600 ms. After a delay, a salient tone of unpredictable pitch sounded in warned trials (as indicated by the musical notation, which was not presented visually, in the second frame); there was silence in the intermingled, and more common, unwarned trials. After a variable, random delay, a visual bar appeared on one side (left in the figure), followed by a bar on the other side at a variable SOA. Both bars then remained until the patient judged verbally whether the left or the right bar had appeared first. In the main experiment, the occasional alerting sounds were always central. In the follow-up study, they were always to the far right of the visual screen.

burst, presented 300–1,100 ms (unpredictably) before the first visual target. The remaining 75% of trials included no tone ('unwarned' trials). Patients were instructed to ignore the occasional sounds, which did not predict the temporal order of left versus right visual events, and were played centrally. After appearing with a particular SOA, the two visual bars remained present until the patient provided an unspeeded 'left first' or 'right first' verbal response. We used standard psychophysical PEST (Parameter Estimation by Sequential Testing) staircases to estimate the threshold of subjective simultaneity for visual events on the two sides. We estimated these thresholds separately for intermingled trials, which either included an alerting tone or were unwarned. Each staircase yielded a threshold for perceiving left and right visual events as synchronous (that is, the SOA at which neither stimulus appeared to lead consistently). In neurologically healthy individuals, this threshold occurs at temporal asynchronies close to zero², indicating no systematic bias in spatial attention. In right-hemisphere patients with left neglect, the threshold is reached only when the left event leads². Our question was whether phasic alerting by the warning tone could overcome this spatial left neglect.

We tested eight patients with right-hemisphere lesions. Each

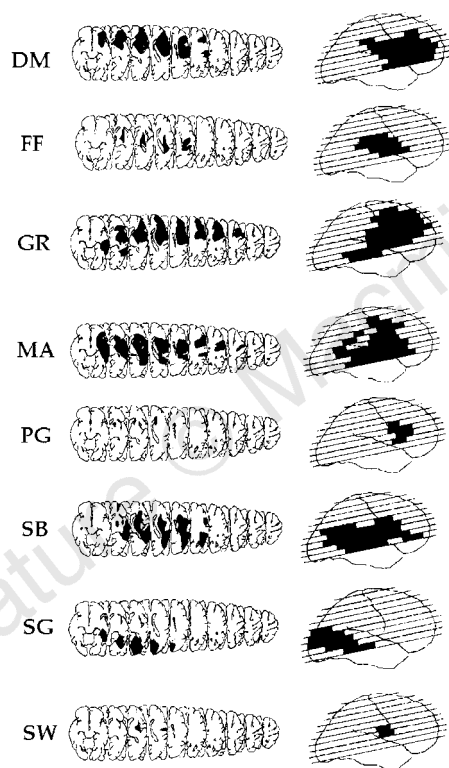


Figure 2 Eight patients with right-hemisphere ischaemic lesions were tested, and were selected if they showed left neglect and a bias of at least 144 ms towards the right in the PEST procedure in our visual task (see Methods) in the unwarned condition (indicating a pathological rightwards bias of spatial attention). All patients had full visual fields on confrontation testing, with the exception of patient S.G. who had a left upper quadrantanopia (in this case, all the peripheral visual stimuli were presented in the lower visual field, below central fixation). Seven patients met the criteria for spatial neglect in the star-cancellation subtest of the behavioural inattention test (BIT)²⁹ (mean left-sided omission score for seven subjects, 13.5, s.d. 13.9). The other patient (S.B.) met the criterion for unilateral inattention in the letter-cancellation subtest of the BIT, omitting 9 of 40 letters. The six males and two females (mean age 64.6 years, s.d. 5.9; mean number of weeks post-stroke, 35.0, s.d. 31.0) all had Computerized Tomography (CT) documented right-hemisphere lesions, which are reconstructed as dark areas on standard templates in the figure²⁶; lesions were all in the right middle cerebral artery territory, except in S.G. who had a posterior lesion affecting the occipitotemporal cortex.

showed left neglect and/or extinction in clinical tests (Fig. 2 legend), and a rightward bias in our visual task when not alerted. Lesions were identified by computed tomography before being plotted onto standard templates²⁶ (Fig. 2). Figure 3 shows how our psychophysical procedure derives thresholds for subjective simultaneity, showing the data for individual staircases from one patient (F.F.). The average thresholds across staircases are plotted for each patient in Fig. 4. For unwarned trials, the average threshold for subjective simultaneity across the eight patients was +497 ms (a substantial left lead). The neglect group thus perceived right visual stimuli as appearing sooner than left visual stimuli, unless the left led by nearly half a second. This confirms that there is a marked delay in the time course of visual awareness for left events relative to right when the patients were not alerted, consistent with their left neglect.

This pathological delay in awareness of left events was changed by presenting an occasional warning sound centrally, to alert the patients phasically. Even though this central sound was uninformative for the visual judgement, and unlateralized, on average it completely abolished the pathological 'prior-entry'^{2,25} of right visual events into the patients' awareness before left events (Fig. 4). The average threshold for subjective simultaneity in warned trials was –7 ms, significantly different from the +497 ms threshold for unwarned trials ($t(7) = 3.70$, $P < 0.01$). We found this change in the threshold for subjective simultaneity when alerted in all eight patients. As warned and unwarned trials were randomly intermingled, and occurred every 2–3 seconds, the effect of the warning tone evidently dissipated rapidly. Thus, the sound could have exerted only a brief phasic effect on alertness, which nevertheless accelerated the perception of left visual events relative to right events as predicted.

An alternative explanation of our results might be that although the central warning sound was located directly between the left and right visual events, its spatial position might still have been advantageous for the perception of left events. As, in the absence of sound, the patients have an attentional bias towards the right, a central sound may effectively draw their attention leftwards (from an initial right position towards the centre). We tested this proposal by presenting the alerting sounds from a loudspeaker hidden to the far right of the visual screen. Under this new configuration, the patients' rightward attentional bias should be maintained or exacerbated if the occasional auditory events (now on the right) merely attract spatial attention towards them. In contrast, if the warning sounds exert their effect through phasic alerting and the associated activation of right-hemisphere networks, the perception of left visual events should again be accelerated relative to right events.

We implemented this follow-up trial with two of our original patients (D.M. and S.B.). In unwarned trials, both patients again showed a strong 'prior-entry' advantage for right visual events, consistent with their left neglect. The pathological temporal disadvantage for left events was again removed in trials in which an auditory warning was included (patient D.M. unwarned threshold +398.0 ms, warned threshold –99.5 ms; S.B., unwarned threshold +306.7 ms, warned threshold –116.6 ms). Thus the effect of the warning tone must be due to phasic alerting, rather than the sound merely attracting attention to its location, as the sound was now on the far right and yet still benefited the left visual event. This is a unique example of left neglect being eliminated by the addition of a salient event on the right. The warning sound not only eliminated the prior-entry advantage for right visual events that was so apparent in unwarned trials: it also led to some advantage for left over right events in the warned trials, thus reversing the prior-entry effect. This reversing pattern was also apparent for four of the eight patients in the original study (Fig. 4).

To test the hemispheric specificity of our results, we repeated the original experiment (that is, with central warning tones) with seven patients with left-hemisphere damage, whose lesions were comparable to those of the right-hemisphere cases. These left-hemisphere patients showed no consistent prior-entry advantage for either side

in the unwarned condition (mean threshold -17.8 ms, s.d. 52.62); unlike the right-hemisphere cases, their performance did not change significantly in warned trials (mean threshold 21.1 ms, s.d. 78.89; $t(6) = 1.09$, not significant).

Several aspects of our methods and results lead us to conclude that, in the right-hemisphere patients with left neglect, phasic alerting by the warning tones specifically accelerated the perceptual processing of left visual events relative to right events, effectively producing a leftwards shift in spatial attention. First, our methods provide a purely perceptual measure of the time course of visual awareness on the left versus right, rather than relying on the speeded manual responses used previously in studies of warning effects on neglect²³, which may have tapped non-perceptual processes, such as motor readiness, instead¹³. Second, our method pits the speed of perceptual processing of left visual events directly against that of right visual events in every trial, rather than measuring performance for each side in isolation. Thus, our results show that phasic alerting affects the spatial pattern of perception in the right-hemisphere patients. Third, not only did the alerting sound eliminate the advantage for right visual events that was apparent in unwarned

trials for these patients but, in many cases, the sound actually reversed this advantage, to produce an advantage for left visual events relative to right events in warned trials. This can only be explained by a leftward shift of spatial attention when phasically alerted. We would not expect to find such reversals on warned trials in all cases. To produce a reversal, the leftward shift in attention caused by phasic alerting must exceed the patient's chronic rightward bias. Our hypothesis only makes the general prediction that some degree of leftward shift should be induced to oppose this bias. Nevertheless, the reversals are striking confirmation that phasic alerting by the sounds did indeed induce leftward attentional shifts, which reached the left of centre in those cases showing reversals.

These results are, to our knowledge, the first direct evidence that phasic alerting of the damaged brain can overcome perceptual spatial neglect in right-hemisphere patients. This is important because, first, it shows that the patients' pathological delay in experiencing left events, in unwarned trials, cannot be due only to afferent losses of sensory input following lesion, as warning sounds cannot restore missing afferent connections in the visual system. Thus, these patients' left-sided deficit is attentional rather than sensory in nature. Second, a warning event in one modality (audition) can dramatically influence the time course of phenomenal experience for an entirely unrelated event in a separate modality (vision). This provides strong neuropsychological support for the claim¹³ that phasic alerting can directly affect the speed of perceptual processing, rather than merely affecting motor readiness. Third, our results concur with physiological and functional imaging data concerning the neural networks involved in alertness and in spatial attention^{3,15,22}. In particular, they agree with the idea of right lateralization of these networks at cortical levels, and, more specifically, with the involvement of right frontal-parietal circuits (damaged in our neglect patients) in tonic alertness, but of ascending subcortical projections (intact in our patients) in phasic alertness^{22,27}. The effect of the warning sounds shows that phasic alerting was preserved in our right-hemisphere patients, despite their tonic impairment.

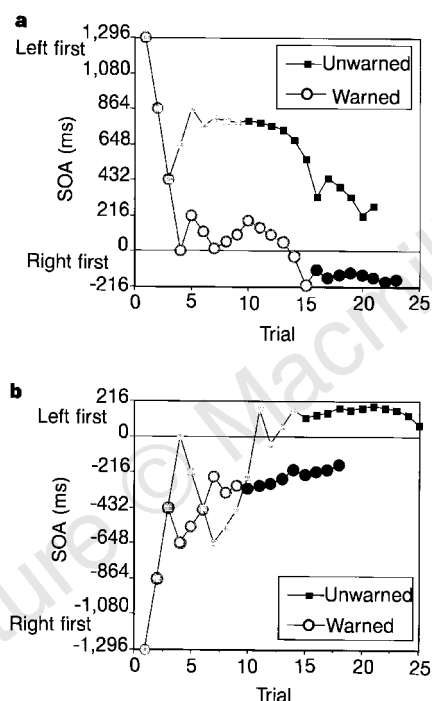


Figure 3 Examples of data from four concurrent staircases for one of the patients (F.F.), showing the PEST staircase procedure which determines thresholds for phenomenal simultaneity in warned (squares) and unwarned (circles) trials. The abscissa shows successive trials from each of the staircases illustrated (in the experiment, successive trials from six concurrent staircases were intermingled in a random order). The ordinate indicates the SOA between the two visual bars that was selected by the staircase for a particular trial. **a**, Performance for one warned staircase and one unwarned staircase, which both began with an extreme left lead (+1,296 ms). **b**, Two staircases that began with an extreme right lead. Each staircase independently adjusted the SOA between the two visual targets, converging on the SOA at which the patient was equally likely to report the left or the right stimulus as appearing first. The stepping rules in the psychophysical staircase followed a standard PEST algorithm²⁸; briefly, this determines the size and direction of successive steps in SOA according to whether the response at a given SOA is reversed from the previous response in the staircase. To determine the threshold for subjective simultaneity, only the data after the fifth response reversal were averaged (that is, only the data points shown in darker shading, towards the right of the figure), so that the staircase had time to settle. Individual data such as those shown here are inevitably somewhat 'noisy' compared with the average results. However, it can be seen that F.F. required a substantial left lead for subjective simultaneity when unwarned, but not when warned.

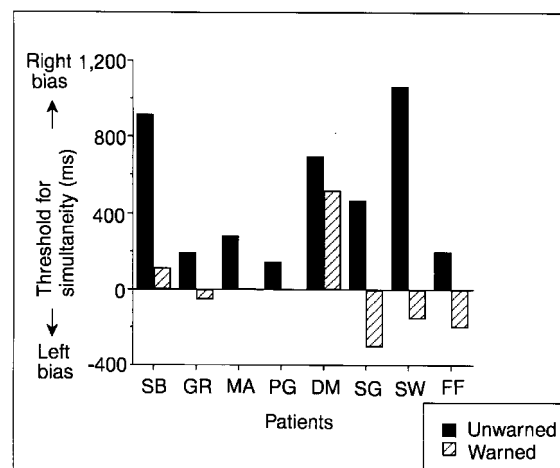


Figure 4 Mean thresholds of subjective simultaneity for each of the eight right-hemisphere-damaged patients, as derived from the PEST staircases they underwent. Positive values indicate that phenomenal synchrony required the left event to appear first (consistent with a rightward bias in attention); negative values indicate that a right lead was required (suggesting leftward attention). In unwarned trials, the average threshold for subjective simultaneity was +497 ms (s.d. 358 ms; range 149–1,069 ms), consistent with severe left neglect. The average threshold for subjective simultaneity in warned trials was only -7.2 ms (s.d. 251 ms; range -304 to $+521$ ms), showing that the left neglect was eliminated. Every patient showed a reduction in the leftward temporal lead required to produce phenomenal simultaneity in warned versus unwarned trials. Moreover, four patients (S.W., S.G., G.R., F.F.) actually required a rightward lead when phasically alerted by the warning sound, a pattern which was also seen in the follow-up study.

Finally, our results indicate that the tonic defect in alertness in right-hemisphere patients contributes to their spatial neglect, because phasically increasing alertness can overcome their spatial deficit in visual awareness. The data show that a nonspatial intervention (that is, the presentation of warning tones, regardless of their location) can have a beneficial spatial effect in right-hemisphere-damaged patients, ameliorating their disabling bias towards the right. This indicates that there may be new approaches to rehabilitation. Most existing treatments of spatial neglect have limited success¹. They typically seek to direct the patient's attention to the impaired left side, either by presenting lateralized visual cues at the left (which has the disadvantage that such cues are often themselves neglected), or by encouraging the patient to concentrate voluntarily on the affected side (this treatment typically shows only limited transfer to unsupervised situations)¹. Our results indicate that any manipulation that phasically increases alertness, such as warning sounds, may effectively treat not just the tonic deficit of alertness in right-hemisphere patients, but also their spatial neglect. □

Methods

Visual stimuli were presented on an active matrix LCD computer monitor (70 Hz refresh rate), at a viewing distance of ~40 cm. All stimuli appeared white against a uniformly black background. At the start of each trial, a fixation cross (0.9° × 0.9°) appeared at the centre of the display. Subjects were repeatedly requested to maintain their gaze on the central cross, and this was monitored in each trial by an observer. Less than 10% of trials were excluded because of fixation failures. Visual targets consisted of two horizontal bars, each subtending 0.4° in height and 3.1° in width. The bars appeared at symmetrical locations in the left and right visual fields, at the same height as the fixation cross (except for patient S.G.; Fig. 2), with their outer edges 9.8° from fixation.

The sequence of events for a warned trial with a left temporal lead is shown in Fig. 1. In all trials, the fixation cross appeared alone for 600 ms. In 25% of trials, a 300-ms tone burst (randomly either 400 Hz (89 dB) or, 1000 Hz (84 dB)) was then presented centrally from two loudspeakers (Macintosh Multimedia Model SK-A10) hidden directly behind the monitor (to the right of the monitor in the follow-up experiment). The use of such warning tones is a standard way to phasically alert neurologically normal people^{6,10,11,13}. At a random delay of 300–1,100 ms after the warning tone ended, the two bars appeared to the left and right of fixation, with a variable SOA between them of 0–1,296 ms. In the remaining 75% of trials, the same sequence of events was used, except that a 300-ms silence replaced the tone. In every trial, subjects made a forced choice regarding which of the two bars appeared first. The bars remained on the screen until this verbal response. Each patient underwent eight separate but concurrent psychophysical staircases in a randomly intermingled sequence of trials (six of the eight staircases were unwarned and two were warned; in all cases, half of the staircases began with an extreme left lead (1,296 ms) and half with a 1,296-ms lead for the right instead). Each staircase independently adapted to the patients' responses, adjusting SOA to the point where left-first responses were as common as right-first responses for that staircase, with the steps in SOA adjusted by the PEST algorithm²⁸. Trials continued until the patient had made at least ten reversals of response in one of the two warned staircases and three of the six unwarned staircases (typically after about 200 trials per subject). Following standard psychophysical practice, and to allow staircases to settle, the threshold for subjective simultaneity was computed from each staircase by averaging SOA in all trials after the fifth response reversal in that staircase.

Selection criteria for the right-hemisphere patients are described in the Fig. 2 legend. Only patients D.M. and S.B. from the original right-hemisphere group could take part in the follow-up study (over a year later); the other cases had by then recovered, or were unavailable or deceased. The seven left-hemisphere-damaged patients (four female, three male) were matched for age with the right-hemisphere-damaged group (mean age 60.1, s.d. 15.5) and for extent of lesion, and were tested an average of 34.4 (s.d. 6.6) days after their stroke.

Received 26 May; accepted 29 June 1998.

- Robertson, I. H. & Marshall, J. C. *Unilateral Neglect: Clinical and Experimental Studies* (Lawrence Erlbaum, Hillsdale, 1993).
- Rorden, C., Mattingley, J. B., Karnath, H. O. & Driver, J. Visual extinction and prior entry: impaired

- perception of temporal order with intact motion perception after unilateral parietal damage. *Neuropsychologia* 35, 421–433 (1997).
- Pardo, J. V., Fox, P. T. & Raichle, M. E. Localization of a human system for sustained attention by positron emission tomography. *Nature* 349, 61–64 (1991).
- Heilman, K. M. & Abell, T. V. D. Right hemisphere dominance for mediating cerebral activation. *Neuropsychologia* 17, 315–321 (1979).
- Wilkins, A. J., Shallice, T. & McCarthy, R. Frontal lesions and sustained attention. *Neuropsychologia* 25, 359–365 (1987).
- FernandezDuque, D. & Posner, M. I. Relating the mechanisms of orienting and alerting. *Neuropsychologia* 35, 477–486 (1997).
- Robertson, I. H., Tegner, R., Tham, K., Lo, A. & Nimmo-Smith, I. Sustained attention training for unilateral neglect: theoretical and rehabilitation implications. *J. Clin. Exp. Neuropsychol.* 17, 416–430 (1995).
- Robertson, I. H. *et al.* Auditory sustained attention is a marker of unilateral spatial neglect. *Neuropsychologia* 35, 1527–1532 (1997).
- Posner, M. I. *Chronometric Explorations of Mind* (Lawrence Erlbaum, Hillsdale, 1978).
- Posner, M. I. & Peterson, S. E. The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42 (1990).
- Parasuraman, R., Warm, J. S. & Dember, W. N. in *Ergonomics and Human Factors: Recent Research* (eds Mark, L., Warm, J. S. & Huston, R. L.) 11–32 (Springer, New York, 1987).
- Mackworth, J. F. Vigilance, arousal and habituation. *Psychol. Rev.* 75, 308–322 (1968).
- Hackley, S. A. & Vallenclan, F. Automatic alerting does not speed late motoric processes in a reaction-time task. *Nature* 391, 786–788 (1998).
- Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Ann. Rev. Neurosci.* 18, 193–221 (1995).
- Corbetta, M., Miezin, F. M., Shulman, G. L. & Peterson, S. E. A PET study of visuospatial attention. *J. Neurosci.* 13, 1202–1206 (1993).
- Mesulam, M. M. A cortical network for directed attention and unilateral neglect. *Ann. Neurol.* 10, 309–325 (1981).
- Kinsbourne, M. in *Unilateral Neglect: Clinical and Experimental Studies* (eds Robertson, I. H. & Marshall, J. C.) 63–86 (Lawrence Erlbaum, Hillsdale, 1993).
- DiPellegrino, G. & De Renzi, E. An experimental study on the nature of extinction. *Neuropsychologia* 33, 153–170 (1995).
- Paus, T. *et al.* Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. *J. Cog. Neurosci.* 9, 392–408 (1997).
- Hockey, G. R. J. Effect of loud noise on attentional selectivity. *Q. J. Exp. Psychol.* 22, 28–36 (1970).
- Rueckert, L. & Graham, J. Sustained attention deficits in patients with right frontal lesions. *Neuropsychologia* 34, 953–963 (1996).
- Heilman, K. M., Schwartz, H. D. & Watson, R. T. Hypoarousal in patients with the neglect syndrome and emotional indifference. *Neurology* 28, 229–232 (1978).
- Posner, M. I., Inhoff, A. W. & Friedrich, F. J. Isolating attentional systems: a cognitive-anatomical analysis. *Psychobiol.* 15, 107–121 (1987).
- Kinomura, S., Larsson, J., Gulyas, B. & Roland, P. E. Activation by attention of the human reticular formation and thalamic intralaminar nuclei. *Science* 271, 512–515 (1996).
- Titchener, E. B. *Lectures on the Elementary Psychology of Feeling and Attention* (Macmillan, New York, 1908).
- Damasio, A. *Lesion Analysis in Neuropsychology* (Oxford Univ. Press, Oxford, 1989).
- Heilman, K. M., Watson, R. T., Valenstein, E. & Goldberg, M. E. in *Handbook of Physiology, Section 1: The Nervous System* (ed. Plum, F.) 461–481 (American Physiological Society, Bethesda, MD, 1987).
- Taylor, M. M. & Creelman, C. D. PEST: efficient estimates on probability functions. *J. Acoust. Soc. America* 41, 782–787 (1967).
- Wilson, B. A., Cockburn, J. & Halligan, P. *Behavioural Inattention Test* (Thames Valley Test Company, Farnham, 1988).

Acknowledgements. This research was supported by the Medical Research Council (UK) and also by a Wellcome Trust grant to J.D. J.B.M. was supported by a National Health and Medical Research Council (Australia) Neil Hamilton Fairley Fellowship, and by AMRAD Australia.

Correspondence and requests for materials should be addressed to I.H.R.

Increased number of synaptic GABA_A receptors underlies potentiation at hippocampal inhibitory synapses

Zoltan Nusser*†, Norbert Hájos†§, Peter Somogyi* & Istvan Mody†

* Medical Research Council, Anatomical Neuropharmacology Unit, Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3TH, UK
† Departments of Neurology and Physiology, UCLA School of Medicine, Los Angeles, California 90095-1769, USA

Changes in synaptic efficacy are essential for neuronal development¹, learning and memory formation² and for pathological states of neuronal excitability, including temporal-lobe epilepsy³. At synapses, where there is a high probability of opening of postsynaptic receptors⁴, all of which are occupied by the

† Present address: Department of Neurology, UCLA School of Medicine, Los Angeles, California 90095-1769, USA.

§ Present address: Institute for Experimental Medicine, Hungarian Academy of Sciences, Budapest, Hungary.

released transmitter⁵⁻⁹, the most effective means of augmenting postsynaptic responses is to increase the number of receptors^{2,10,11}. Here we combine quantal analysis of evoked inhibitory postsynaptic currents with quantitative immunogold localization of synaptic GABA_A receptors in hippocampal granule cells in order to clarify the basis of inhibitory synaptic plasticity induced by an experimental model of temporal-lobe epilepsy (a process known as kindling)¹⁰. We find that the larger amplitude (66% increase) of elementary synaptic currents (quantal size) after kindling results directly from a 75% increase in the number of GABA_A receptors at inhibitory synapses on somata and axon initial segments. Receptor density was up by 34–40% and the synaptic junctional area was expanded by 31%. Presynaptic boutons were enlarged, which may account for the 39% decrease in the average number of released transmitter packets (quantal content). Our findings establish the postsynaptic insertion of new GABA_A receptors and the corresponding increase in postsynaptic responses augmenting the efficacy of mammalian inhibitory synapses.

At excitatory and inhibitory synapses of the adult mammalian

central nervous system (CNS) the gain of synaptic transmission can undergo lasting and marked alterations. The plasticity of excitatory (glutamatergic) synapses has been widely studied, but no consensus exists regarding the pre- or postsynaptic site of the alteration responsible for the change in synaptic gain^{2,11}. Far fewer studies have addressed the long-lasting plasticity at central inhibitory (GABAergic) synapses^{10,12-15}. A change in the size of miniature inhibitory postsynaptic currents (mIPSCs) occurs at GABAergic synapses on granule cells of the dentate gyrus after kindling-induced epilepsy, and an increased number of synaptic GABA_A receptors was inferred from non-stationary noise analysis¹⁰. Here we have used a direct approach to determine the relationship between the size of synaptic responses and the number of receptors by combining physiological recordings with electron microscopy^{9,13}.

For quantal analysis, IPSCs were evoked by local stimulation of inhibitory axons, which very probably terminate perisomatically on granule cells. Analogously to earlier observations in young animals⁵, this type of stimulation in adult rats evoked IPSCs with amplitude distributions having clearly distinguishable multiple peaks (Fig. 1). The peaks were fitted with multiple gaussian distributions, in each case revealing equal (quantal) spacings between successive peaks. As generally noted at CNS synapses (see, for example, ref. 5), larger peaks fit less well, and the quantal variance fails to increase linearly, but according to an autocorrelation analysis¹⁶ performed in four cells (including the two shown in Fig. 1), the likelihood for the chance occurrence of the observed peaks in the distributions was $< 4 \times 10^{-5}$. The quantal analysis was done in accordance with published methods of Edwards *et al.* (ref. 5 and Methods), and revealed a significant 66% increase in the quantal size (q) of IPSCs in kindled neurons (Table 1). The quantal content (m) obtained by the method of failures [$m_f = \ln(N/N_0)$] assuming Poisson release statistics did not correspond to the value calculated directly (m_d) by dividing the average evoked IPSC (eIPSC) by q .

Considering full occupancy of postsynaptic GABA_A receptors by the released transmitter^{5,6}, and a large open probability ($P_o = 0.8$) at the peak of synaptic currents⁴, our findings indicate that control synapses have on average 26 GABA_A receptors ($= 435 \text{ pS}/(21 \text{ pS} \times 0.8)$) considering an average conductance of 21 pS, and this number increases to 43 after kindling. Alternatively, the 66% increase in q may result from (1) an increase in P_o or (2) an augmented single-channel conductance. (1) An increase in P_o to its maximal possible value of 1.0 can account for no more than 25% ($1.0/0.8 = 1.25$) of the augmented q , but our experimental results are consistent with a decrease, if any change, in P_o after kindling (see below). (2) An enhanced contribution of large conductance states to the average conductance of synaptic GABA_A receptors can also be excluded. Peak-scaled non-stationary fluctuation analysis^{6,8,10} of eIPSCs revealed similar mean conductances of GABA_A receptors in control ($20.6 \pm 0.8 \text{ pS}$; $n = 5$) and kindled ($21.3 \pm 1.4 \text{ pS}$; $n = 4$) granule cells.

To examine the effect of kindling on the expression of GABA_A receptor subunits in the dentate gyrus, light-microscopic immunocytochemistry was performed for the $\alpha 1$, $\alpha 2$, $\beta 2/3$ and $\gamma 2$ subunits. An increased immunostaining of granule cells was observed for all subunits tested ($n = 5$ pairs of animals), which was apparent from the enhanced staining of the granule cell layer and the neuropile of the molecular layer. No immunostained axon terminals could be detected around somata and axon initial segments (AISs) of granule cells, indicating that all of the detectable change in the expression of GABA_A receptor subunits takes place postsynaptically.

To investigate directly the relation between the increased quantal size and the number of synaptic receptors, we applied postembedding immunogold localization of the $\beta 2$ and $\beta 3$ subunits. The immunosignal was targeted against these subunits because: (1) they are the major β subunits expressed by dentate granule cells¹⁷; (2) they are expressed at perisomatic synapses¹⁸ and (3) every functional GABA_A receptor probably contains β subunits¹⁹. Thus, alterations

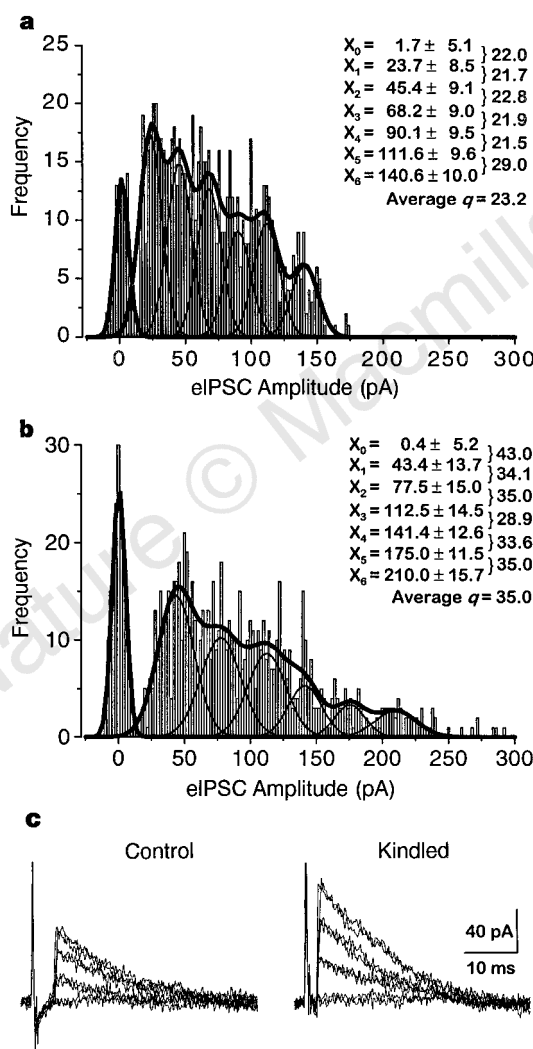


Figure 1 Increased quantal size of eIPSCs in kindled granule cells. A similar number of events (a) control ($n = 780$) and (b) kindled ($n = 782$) were binned at 2 pA. The distributions including failures were fitted by 7 gaussians. Their means $\pm$ s.d. (baseline variance subtracted) are indicated, and quantal size was calculated as the average difference between consecutive peaks. Baseline noise was comparable in both cells (control, $-0.2 \pm 2.6 \text{ pA}$; kindled, $0.1 \pm 2.9 \text{ pA}$). c, Representative traces of failures and eIPSCs selected from the first three peaks of the amplitude distributions.

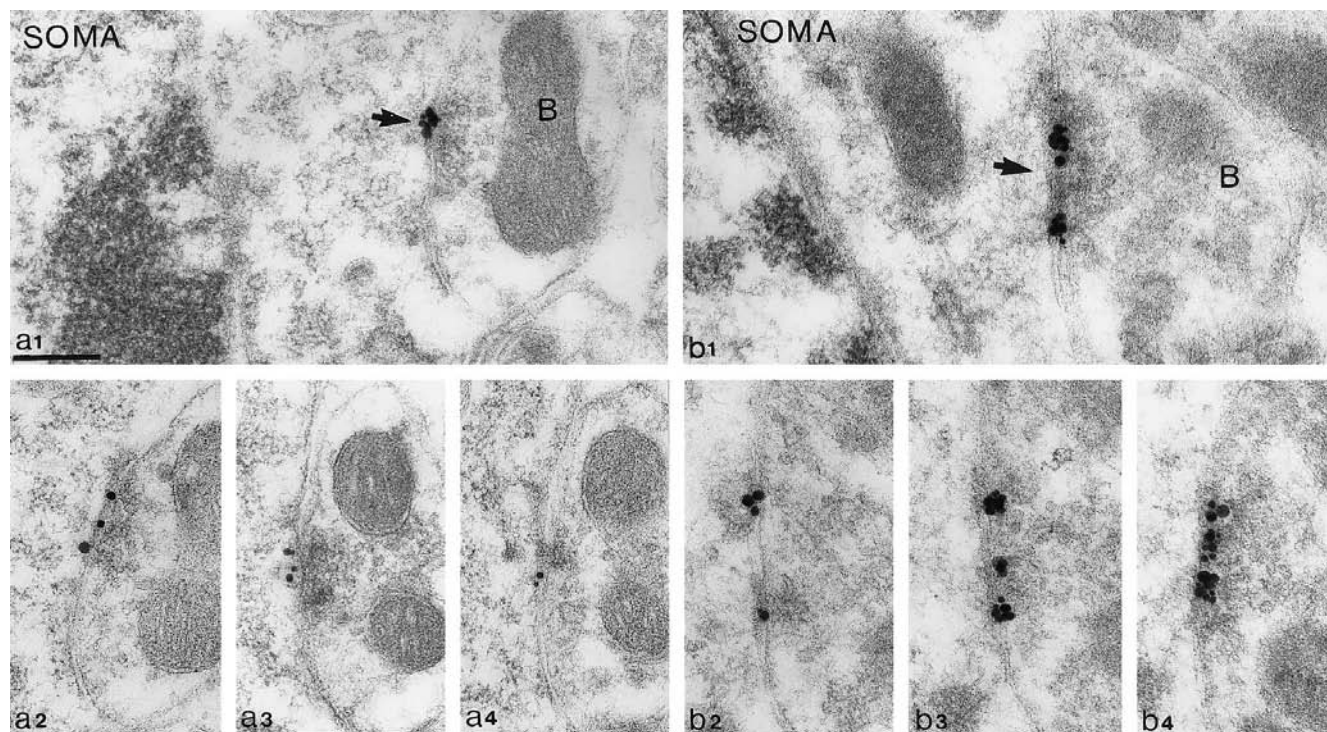


Figure 2 GABA_A receptor (β2/3 subunits) immunoreactivity in somatic synapses of granule cells from rats. **a**, Control; and **b**, kindled animals. Immunopositive synapses (arrows) made by synaptic boutons (B) contain clusters of immunoparticles. The whole synapse on a control granule cell (**a**₁–**a**₄) is shown and

contains 13 immunoparticles. The synapse on a kindled granule cell was present in six serial sections (four shown: **b**₁–**b**₄) and had a total of 71 particles. The full area of the synapse in **a** is smaller (0.040 μm²) than that in **b** (0.111 μm²). Scale bar, 0.2 μm.

in the total number of GABA_A receptors at synapses should be faithfully reflected by changes in the number of β2/3 subunits. As our evoked responses are of proximal origin and mIPSCs are likely to originate from perisomatic synapses²², here we examined the β2/3 subunit content of somatic and AIS synapses only. To determine the immunoreactive receptor content of the whole synapse, the entire postsynaptic area was reconstructed from serial sections (for example Fig. 2a₁–a₃; ref. 9). After kindling, the immunoreactive β2/3 subunit content of synapses was increased on both somata (Fig. 2) and AISs (Fig. 3) of granule cells, with a somewhat larger increase at somatic synapses (Fig. 4a; soma, 94%; AIS, 61%). The larger number of immunoparticles was the consequence of a 34% and 40% increase in their density at somatic and AIS synapses, respectively (Fig. 4a), as well as an average 31% increase in the area (Fig. 4a) of synaptic junctions pooled from both compartments. To exclude the possibility that the increase at synapses was restricted to β2/3 subunits, and to establish whether some of the increased expression observed at the light microscopic level for the other subunits was also localized to synapses, we determined the quantitative distribution of α2 subunits at AIS synapses²¹. All synapses on AISs were found to be consistently immunopositive for α2 subunits (control, 6.4 ± 4.2 particle per synapse; range 1–17; *n* = 14), and kindling (10.1 ± 6.2; range 6–31; *n* = 15) increased their immunoreactive receptor content by 58%. The remarkable similarity between the increase in the number of α2 (58%) and β2/3 (61%) subunits at AIS synapses suggests that alterations detected with subunit-selective

antibodies represent changes in the total number of synaptic GABA_A receptors.

We next examined how the observed increase in the number of GABA_A receptors at perisomatic synapses is correlated with the increased *q*. The immunogold data were obtained from a large number of perisomatic synapses (*n* = 84 from three control rats and *n* = 98 from three kindled rats) of several postsynaptic neurons. In contrast, a given *q* was measured at a few synapses made by a single presynaptic axon²². Therefore, we initially sought to compare the distributions of immunoparticles (Fig. 4b) with those of mIPSCs (Fig. 4c) pooled from several postsynaptic cells. A large variability in the immunoreactive GABA_A receptor content was observed in both subcellular compartments of control granule cells (soma, mean = 8.9 ± 0.9 particles per synapse, range 0–30, c.v. = 0.75 ± 0.05; AIS, mean = 13.6 ± 3.4 particles per synapse, range 2–35, c.v. = 0.58 ± 0.12), resulting in a skewed distribution. The large variability and the skewed distribution persisted after kindling (soma, c.v. = 0.78 ± 0.12; AIS, c.v. = 0.62 ± 0.05) with an approximately parallel shift in the cumulative probability plot (Fig. 4b). A similar shift was observed in the cumulative distribution of mIPSC conductance (mIPSG) recorded in kindled granule cells (Fig. 4c). The relative increase in the median receptor number (75%) after kindling was larger than that observed for mIPSG (41%), possibly due to changes at some non-overlapping synapse populations sampled for mIPSCs and ultrastructural analysis.

In contrast to mIPSCs, there was a good match between the

Table 1 Quantal parameters of inhibitory postsynaptic currents originating from the perisomatic region of control and kindled granule cells

Current	<i>n</i>	<i>N</i> – <i>N</i> ₀	<i>N</i> ₀	<i>q</i> (pS)	<i>m</i> _d	<i>m</i> _f	c.v. (%)
Control eIPSCs	8	467 ± 60	142 ± 19	435 ± 23	2.6 ± 0.3	1.5 ± 0.1	26 ± 2
Kindled eIPSCs	10	487 ± 64	192 ± 25	722 ± 71*	1.6 ± 0.2*	1.3 ± 0.1	30 ± 3
Control mIPSCs	4	264 ± 66	–	436 ± 17	–	–	34 ± 3
Kindled mIPSCs	4	426 ± 89	–	615 ± 17*	–	–	40 ± 2

N – *N*₀, number of events per cell, where *N* is the total number of responses and *N*₀ is the number of failures; *q*, average quantal size; *m*_d, quantal content (direct method); *m*_f, quantal content (method of failures); c.v., coefficient of variation (after subtraction of baseline noise variance). The *q* of mIPSCs is given as the median of cumulative probability distributions.

* Significant difference from control (*P* < 0.05, ANOVA). Data are expressed as mean ± s.e.m.

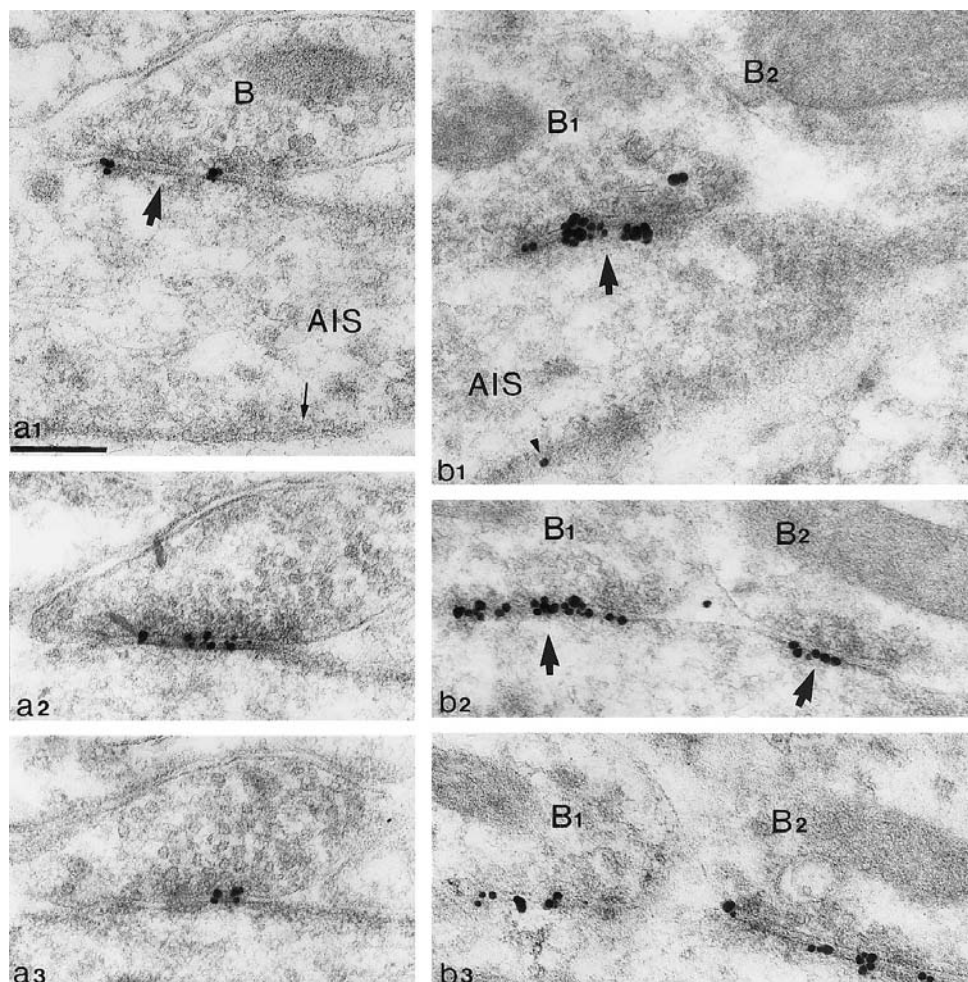


Figure 3 Kindling increases the immunoreactive $\beta 2/3$ subunit content of AIS synapses. **a**, Control; **b**, kindled AISs were identified by membrane undercoating (small arrow). Immunoparticles are concentrated at synaptic junctions (arrows) made by boutons of axon cells (B) and are occasionally present on extrasynaptic membranes (arrowhead). The area of the synapse in **a** is

$\sim 0.060 \mu\text{m}^2$ and contains 18 particles. Two synapses, converging onto an AIS of a kindled granule cell (**b**), contain 75 (B_1 , area $\approx 0.131 \mu\text{m}^2$) and 19 particles (B_2 , area $\approx 0.060 \mu\text{m}^2$). Note the presence of more receptor clusters in kindled (**b**) compared with control (**a**) rats. Scale bar, $0.2 \mu\text{m}$.

change in q of eIPSCs and the increase in the number of immunolabelled receptors. This direct measurement also allowed us to derive possible alterations in the P_o of the postsynaptic receptors after kindling. According to our electrophysiological measurements, 26 functional GABA_A receptors should be present at control synapses where quantal eIPSCs are generated. The weighted mean (64% somatic and 36% AIS synapses, to reflect their natural proportion on granule cells²³) of gold particles was 10.2 for $\beta 2/3$ subunits, corresponding to ~ 2.5 functional GABA_A receptors per particle. On the basis of this relationship, the 18.5 particles per synapse after kindling should correspond to ~ 47 functional receptors, which is within 10% of the predicted 43 channels derived from the measured q when $P_o = 0.8$. Thus, regardless of the initial estimate of P_o at control synapses, the increase in q can be solely attributed to the more receptors present without an accompanying increase in P_o . As more receptors are inserted into synaptic membranes after kindling, and as the size of the synapses increases, at some synapses the receptors may not be fully occupied by GABA. It has been reported⁹ that postsynaptic GABA_A receptors have a very high degree of occupancy at synapses containing less than ~ 80 receptors (corresponding to ~ 33 particles in our study). Because 99% of the perisomatic synapses in control granule cells contain < 33 particles, full receptor occupancy^{5,6} probably occurs at most of these synapses. In contrast, $\sim 10\%$ of the kindled synapses contain > 33 particles, which is consistent with an incomplete occupancy of

the receptors, whereupon the size of IPSCs at such synapses will depend on presynaptic factors, such as the transmitter concentration in the cleft.

We have found both anatomical and physiological evidence for presynaptic changes after kindling. With the postsynaptic insertion of GABA_A receptors and the enlargement of the postsynaptic area, the size of the presynaptic boutons increased as well (sectioned bouton area in control, $0.20 \pm 0.02 \mu\text{m}^2$; kindling, $0.29 \pm 0.01 \mu\text{m}^2$; $P > 0.01$). Larger boutons could decrease the efficacy of action potentials for evoking release, which might explain the 39% decrease in quantal content observed after kindling. Alternative possibilities include changes in the density of presynaptic GABA_A, metabotropic glutamate, kainate receptors, or Ca^{2+} and K^+ channels. It is also interesting to note that in spite of a sufficiently large postsynaptic space available for the insertion of new GABA_A receptors^{9,24}, the postsynaptic area had to increase to accommodate the new receptors. If one accepts that GABA_A receptors form small clusters in the postsynaptic membrane^{9,24}, our findings are consistent with the necessity for a physical separation between such clusters.

In summary, our findings demonstrate a direct relationship between synaptic GABA_A receptor number and quantal size at potentiated GABAergic synapses in the adult brain. The kindling-induced change in synaptic efficacy also includes an enlargement of the synaptic area accompanied by increased presynaptic bouton size

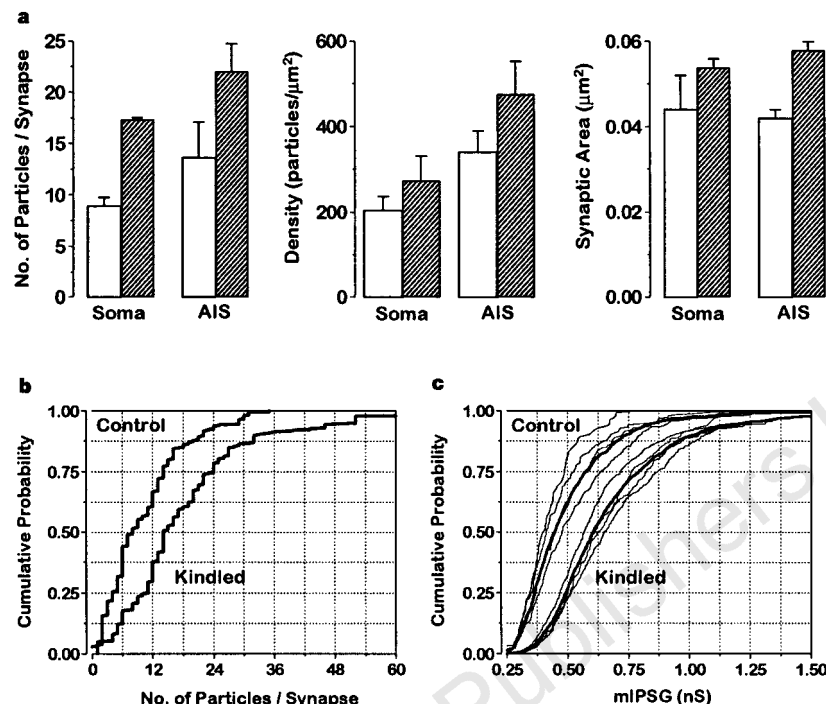


Figure 4 Distribution of synaptic parameters. **a**, The average number of immunoparticles labelling synaptic $\beta 2/3$ subunits was significantly larger (Student's t -test $P < 0.05$) in kindled ($n = 3$; shaded bars) than in control rats ($n = 3$). There was an increased immunoparticle density at somatic (control, 204 ± 33 particles per μm^2 ; kindled, 274 ± 57 particles per μm^2) and AIS (control, 340 ± 50 particles per μm^2 ; kindled, 476 ± 78 particles per μm^2) synapses, and an enlargement of the somatic (control, $0.044 \pm 0.008 \mu\text{m}^2$; kindling, $0.054 \pm 0.002 \mu\text{m}^2$) and AIS (control, $0.042 \pm 0.002 \mu\text{m}^2$; kindled, $0.058 \pm 0.002 \mu\text{m}^2$) synaptic junctional area. **b**, Cumulative weighted distributions

of immunoparticles at AIS and somatic synapses in control ($n = 301$) and kindled ($n = 287$) preparations (pooled from three animals each). **c**, Ensemble distributions of mIPSG (thick lines) recorded in control and kindled granule cells ($n = 4$ each, thin lines). The means $\pm$ s.d. of the ensemble distributions were 496 ± 195 and 678 ± 288 pS in control and kindled preparations respectively. Owing to differences in recording temperature, intracellular solutions and selection criteria, these values are somewhat lower than those previously published¹⁰, but in agreement with that study, no differences were found in the kinetics of mIPSCs of control and kindled neurons.

and decreased quantal content in agreement with the 'synaptic homeostasis' principle suggested to occur at *Drosophila* neuromuscular junctions²⁵. At some large kindled synapses the receptors might lose full occupancy, whereupon further insertion of GABA_A receptors will no longer linearly increase postsynaptic responses. Larger inhibitory events in epilepsy might be puzzling, but such events might better synchronize²⁶ granule cell firing. Alternatively, many of the newly inserted GABA_A receptors might not share the molecular makeup of the existing channels; for example their increased Zn^{2+} sensitivity might lead to the collapse of inhibition at a critical point during seizures²⁷. □

Methods

Preparation of animals and tissue. The procedure of daily kindling of Wistar rats through the hippocampal commissures (100 μA , 60 Hz for 1 s) has been described earlier^{10,27}. Control (implanted but never stimulated, $n = 5$) and kindled animals (with 43–49 seizures, $n = 5$) were perfused 48 h after the last seizures with a fixative containing 0.05% glutaraldehyde, 4% paraformaldehyde and $\sim 0.2\%$ picric acid in 0.1 M phosphate buffer for 13–30 min. The brains were then removed, 500- μm -thick Vibratome sections from the dorsal hippocampus were cut and were embedded into an acrylic resin at -50°C (Lowycriol HM20, Chemische Werke Lowi GmbH, Germany)^{9,18}.

Antibodies and controls. The characterization of polyclonal antibodies to the $\alpha 1$ (ref. 28), $\alpha 2$ and $\gamma 2$ subunits and the monoclonal antibody (code no. bd17) against the $\beta 2$ and $\beta 3$ subunits²⁹, has been described earlier. Antibodies were used at the following final protein concentrations (in $\mu\text{g mL}^{-1}$): $\alpha 1$, 1.3; $\alpha 2$, 1 and 15 for pre- and postembedding reactions, respectively; $\beta 2/3$, 40 and 120 for pre- and postembedding reactions, respectively; $\gamma 2$, 1. Selective labelling, resembling that obtained with the specific antibody, could not be detected when the primary antibody was either omitted or replaced by 5% normal serum.

Pre-embedding immunocytochemistry. A standard procedure¹⁸ was applied on 70- μm -thick Vibratome sections using biotinylated secondary antibodies and avidin biotinylated horseradish peroxidase complex.

Postembedding immunocytochemistry. Serial sections (20–25 per grid; 70 nm thick) were cut from the upper blade of the dorsal dentate gyrus and were picked up on polyloform-coated slot grids. The sections were incubated in blocking solutions for 30 min, followed by an incubation with the primary and secondary antibodies^{9,18}. Quantification of immunoreactivity for GABA_A receptor subunits was performed as described earlier⁹. Briefly, the granule cell layer and the hilus of the dentate gyrus were systematically searched until a symmetrical synapse was found between an axon terminal and the soma or an AIS of a granule cell. The synapse was photographed, followed in serial sections completely through the synaptic specialization, which was measured. Immunoparticles were counted within the membrane specialization. The synaptic area was calculated by multiplying its length in each section by the average thickness of sections (~ 70 nm). Tangentially sectioned synapses were excluded from the analysis of synaptic area and immunoparticle density. The sectioned area of axon terminals was measured with a digitizing table at the level where the longest synaptic specialization was observed. No correction for shrinkage or expansion was made.

Electrophysiology. Horizontal 400- μm -thick brain slices were prepared¹⁰ 36–48 h after the last seizure (average number of seizures = 65). Whole-cell patch clamp recordings were obtained at room temperature ($22\text{--}23^\circ\text{C}$) from dentate granule cells visualized by infrared DIC (differential interference contrast) videomicroscopy (Zeiss Axioscope)³⁰. Slices were perfused with an artificial cerebrospinal fluid³⁰ containing 2 mM kynurenic acid and 1 μM tetrodotoxin (TTX). For minimal stimulation experiments, the extracellular $\text{Ca}^{2+}/\text{Mg}^{2+}$ ratio was changed to 1:2 (1.5 mM CaCl_2 and 3 mM MgCl_2)⁵ and TTX was omitted. The intracellular solution contained (in mM) 130 Cs-gluconate, 2 CsCl, 2 MgCl_2 , 20 HEPES, 2 ATP and 1% biocytin (pH 7.25 adjusted with CsOH; osmolality 300–310 mosM). Minimal stimulation (20 μs pulse width) was

delivered through a theta glass pipette (2 mm diameter, pulled to 1 μ m tip diameter) filled with 1 M NaCl and placed 15–30 μ m away from the somata. Stimulation at 0.2–1 Hz caused no paired-pulse depression or facilitation. Recordings³⁰ were done at holding potentials of +10 $\pm$ 5 mV, and the IPSC reversal potential was always determined. Access resistances (compensated 70–80%) were frequently monitored and remained constant ($\pm$ 15%) during the recordings (control, 12.1 $\pm$ 0.5 M Ω ; kindled, 12.2 $\pm$ 0.5 M Ω). Signals were filtered at 2 kHz³⁰, digitized at 6 kHz and analysed with CDR or SCAN software (courtesy of J. Dempster). For the quantal analysis, a Simplex-based least-squares criterion was used to fit non-constrained gaussian distributions to the eIPSC amplitude histograms⁵, in which peaks could be clearly detected by eye. No more than six gaussian distributions were fitted, as the number of very large compound events was very small to be described accurately by a normal curve. Simple binomial distributions did not describe adequately the frequencies of failures and peaks in the distributions of eIPSCs. In the absence of any morphological data on the number of release sites, we have refrained from applying a compound binomial analysis. At the recorded holding potentials, the peak of the first distribution was always larger than 1.5 times the variance of the baseline noise (σ_N^2). Quantal amplitudes were calculated as described by Edwards *et al.*⁵, and were converted to quantal conductances to eliminate driving force bias. The c.v. was calculated from the mean of the first quantal peak of eIPSC distributions, or the mean of mIPSCs, divided by the square root of the respective σ_N^2 subtracted variances. Non-stationary noise analysis was performed as described^{6,10}.

Received 20 April; accepted 22 June 1998.

1. Singer, W. Development and plasticity of cortical processing architectures. *Science* **270**, 758–764 (1995).
2. Bliss, T. V. P. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
3. Ben-Ari, Y. & Represa, A. Brief seizure episodes induce long-term potentiation and mossy fiber sprouting in the hippocampus. *Trends Neurosci.* **13**, 312–318 (1990).
4. Auger, C. & Marty, A. Heterogeneity of functional synaptic parameters among single release sites. *Neuron* **19**, 139–150 (1997).
5. Edwards, F. A., Konnerth, A. & Sakmann, B. Quantal analysis of inhibitory synaptic transmission in the dentate gyrus of rat hippocampal slices: a patch-clamp study. *J. Physiol. (Lond.)* **430**, 213–249 (1990).
6. De Koninck, Y. & Mody, I. Noise analysis of miniature IPSCs in adult rat brain slices: properties and modulation of synaptic GABA_A receptor channels. *J. Neurophysiol.* **71**, 1318–1335 (1994).
7. Stratford, K. J., Tarczy-Hornoch, K., Martin, K. A. C., Bannister, N. J. & Jack, J. B. Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature* **382**, 258–261 (1996).
8. Silver, R. A., Cull-Candy, S. G. & Takahashi, T. Non-NMDA glutamate receptor occupancy and open probability at a rat cerebellar synapse with single and multiple release sites. *J. Physiol. (Lond.)* **494**, 231–250 (1996).
9. Nusser, Z., Cull-Candy, S. G. & Farrant, M. Differences in synaptic GABA_A receptor number underlie variation in GABA mini amplitude. *Neuron* **19**, 697–709 (1997).
10. Otis, T. S., De Koninck, Y. & Mody, I. Lasting potentiation of inhibition is associated with an increased number of γ -aminobutyric acid type A receptors activated during miniature inhibitory postsynaptic currents. *Proc. Natl Acad. Sci. USA* **91**, 7698–7702 (1994).
11. Malenka, R. C. & Nicoll, R. A. Silent synapses speak up. *Neuron* **19**, 473–476 (1997).
12. Komatsu, Y. Age-dependent long-term potentiation of inhibitory synaptic transmission in rat visual cortex. *J. Neurosci.* **14**, 6488–6499 (1994).
13. Wan, Q. *et al.* Recruitment of functional GABA_A receptors to postsynaptic domains by insulin. *Nature* **388**, 686–690 (1997).
14. Brussaard, A. B. *et al.* Plasticity in fast synaptic inhibition of adult oxytocin neurons caused by switch in GABA_A receptor subunit expression. *Neuron* **19**, 1103–1114 (1997).
15. Korn, H., Oda, Y. & Faber, D. S. Long-term potentiation of inhibitory circuits and synapses in the central nervous system. *Proc. Natl Acad. Sci. USA* **89**, 440–443 (1992).
16. Stratford, K. J., Jack, J. B. & Larkman, A. U. Calibration of an autocorrelation-based method for determining amplitude histogram reliability and quantal size. *J. Physiol. (Lond.)* **505**, 425–442 (1997).
17. Persohn, E., Malherbe, P. & Richards, J. G. Comparative molecular neuroanatomy of cloned GABA_A receptor subunits in the rat CNS. *J. Comp. Neurol.* **326**, 193–216 (1992).
18. Nusser, Z. *et al.* Immunocytochemical localization of the α 1 and β 2/3 subunits of the GABA_A receptor in relation to specific GABAergic synapses in the dentate gyrus. *Eur. J. Neurosci.* **7**, 630–646 (1995).
19. Sieghart, W. Structure and pharmacology of γ -aminobutyric acid_A receptor subtypes. *Pharmacol. Rev.* **47**, 181–234 (1995).
20. Soltesz, I., Smetters, D. K. & Mody, I. Tonic inhibition originates from synapses close to the soma. *Neuron* **14**, 1273–1283 (1995).
21. Nusser, Z., Sieghart, W., Benke, D., Fritschy, J.-M. & Somogyi, P. Differential synaptic localization of two major γ -aminobutyric acid type A receptor α subunits on hippocampal pyramidal cells. *Proc. Natl Acad. Sci. USA* **93**, 11939–11944 (1996).
22. Buhl, E. H., Halasy, K. & Somogyi, P. Diverse sources of hippocampal unitary inhibitory postsynaptic potentials and the number of synaptic release sites. *Nature* **368**, 823–828 (1994).
23. Kosaka, T. Synapses in the granule cell layer of the rat dentate gyrus: serial-sectioning study. *Exp. Brain Res.* **112**, 237–243 (1996).
24. Mody, I., De Koninck, Y., Otis, T. S. & Soltesz, I. Bridging the cleft at GABA synapses in the brain. *Trends Neurosci.* **17**, 517–525 (1994).
25. Davis, G. W. & Goodman, C. S. Synapse-specific control of synaptic efficacy at the terminals of a single neuron. *Nature* **392**, 82–86 (1998).
26. Cobb, S. R., Buhl, E. H., Halasy, K., Paulsen, O. & Somogyi, P. Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons. *Nature* **378**, 75–78 (1995).
27. Buhl, E. H., Otis, T. S. & Mody, I. Zinc-induced collapse of augmented inhibition by GABA in a temporal lobe epilepsy model. *Science* **271**, 369–373 (1996).
28. Zezula, J., Fuchs, K. & Sieghart, W. Separation of α 1, α 2 and α 3 subunits of the GABA_A-

benzodiazepine receptor complex by immunoaffinity chromatography. *Brain Res.* **563**, 325–328 (1991).

29. Fritschy, J.-M. & Mohler, H. GABA_A-receptor heterogeneity in the adult rat brain: differential regional and cellular distribution of seven major subunits. *J. Comp. Neurol.* **359**, 154–194 (1995).
30. Hájos, N. & Mody, I. Synaptic communication among hippocampal interneurons: properties of spontaneous IPSCs in morphologically identified cells. *J. Neurosci.* **17**, 8427–8442 (1997).

Acknowledgements. We thank Zahida Ahmad and Brian Oyama for excellent technical assistance, Paul Jays for assistance with the photography, Jean-Marc Fritschy for kindly providing the α 2, β 2/3 and γ 2 subunit-selective antibodies, and Werner Sieghart for the α 1 subunit-selective antibody. This study was supported by the Medical Research Council (UK), a European Commission Shared Cost RTD Programme Grant (no. BIO4CT96-0585) and NIH grant NS36142.

Correspondence and requests for materials should be addressed to Z.N. or I.M. (e-mail: mody@ucla.edu).

Cloning of *inv*, a gene that controls left/right asymmetry and kidney development

Toshio Mochizuki*, Yukio Saijoh†‡, Ken Tsuchiya*, Yasuaki Shirayoshi§, Setsuo Takail, Choji Taya†, Hiromichi Yonekawa†, Kiyomi Yamada||, Hiroshi Nihei*, Norio Nakatsuji§, Paul A. Overbeek#, Hiroshi Hamada†‡ & Takahiko Yokoyama†‡

* Department of Medicine, Kidney Center, and † Department of Anatomy and Developmental Biology, Tokyo Women's Medical University, School of Medicine, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8600 Japan

† Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565, Japan

‡ Crest, Japan Science and Technology Corporation, Tokyo 170-0013, Japan

§ Mammalian Development Laboratory, National Institute of Genetics, 1111 Yata, Mishima-shi 411, Japan

|| Department of Genetics, Research Institute, International Medical Center of Japan, 1-21-1 Toyoma, Shinjuku-ku, Tokyo 162, Japan

Department of Laboratory Animal Science, The Tokyo Metropolitan Institute of Medical Science, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 113, Japan

Most vertebrate internal organs show a distinctive left/right asymmetry. The *inv* (inversion of embryonic turning) mutation in mice was created previously by random insertional mutagenesis¹; it produces both a constant reversal of left/right polarity (situs inversus) and cyst formation in the kidneys². Asymmetric expression patterns of the genes *nodal* and *lefty* are reversed in the *inv* mutant^{3–6}, indicating that *inv* may act early in left/right determination. Here we identify a new gene located at the *inv* locus. The encoded protein contains 15 consecutive repeats of an Ank/Swi6 motif^{7,8} at its amino terminus. Expression of the gene is the highest in the kidneys and liver among adult tissues, and is seen in presomite-stage embryos. Analysis of the transgenic genome and the structure of the candidate gene indicate that the candidate gene is the only gene that is disrupted in *inv* mutants. Transgenic introduction of a minigene encoding the candidate protein restores normal left/right asymmetry and kidney development in the *inv* mutant, confirming the identity of the candidate gene.

Organ primordia of vertebrae are symmetrical at first, and then acquire distinctive left/right asymmetry. In mice, embryonic turning and heart looping are the earliest manifestations of left/right asymmetry. Before morphological asymmetry develops, several genes, including *nodal* and *lefty* in mice, are expressed asymmetrically^{3–6,9,10}. Asymmetric expression of *nodal* and *lefty* is reversed^{3–6} in the *inv* mutant. The *inv* mutation produces a constant reversal of the left/right asymmetry, contrasting with several experimental and genetic mutations that randomize alterations in left/right asymmetry^{11–14}. In addition, the mutation accompanies cyst formation in the kidney and jaundice².

The *inv* mutation was found in a family of transgenic mice into which the tyrosinase minigene had been introduced¹², and is thought to be created by insertional mutagenesis². We previously

identified single-copy probes (named p3.2H and p2.3H) that flank the integrated tyrosinase minigene^{2,15}. Analysis of these probes showed that a region next to p3.2H had been deleted whereas a region containing p2.3H had become duplicated (Fig. 1A), and also indicated that intrachromosomal inversion might have occurred¹⁵.

To characterize the *inv* locus further, we screened the ICRF yeast artificial chromosome (YAC) library¹⁶ by hybridization with the flanking probe p3.2H (ref. 2). We identified three positive clones, ICRFFy902E0682, ICRFFy902G0572 and ICRFFy902G022 (abbreviated as E0682, G0571 and G022) (Fig. 1B). Cosmid libraries were constructed from these YACs. We identified two cosmids, E121 and G16, that mapped to opposite sides of the transgenic insert (Fig. 1B). The distance between the two cosmids is about 400 kilobases (kb) in the wild-type genome (Fig. 1B). An inversion, if present, should make one of these clones map onto different regions

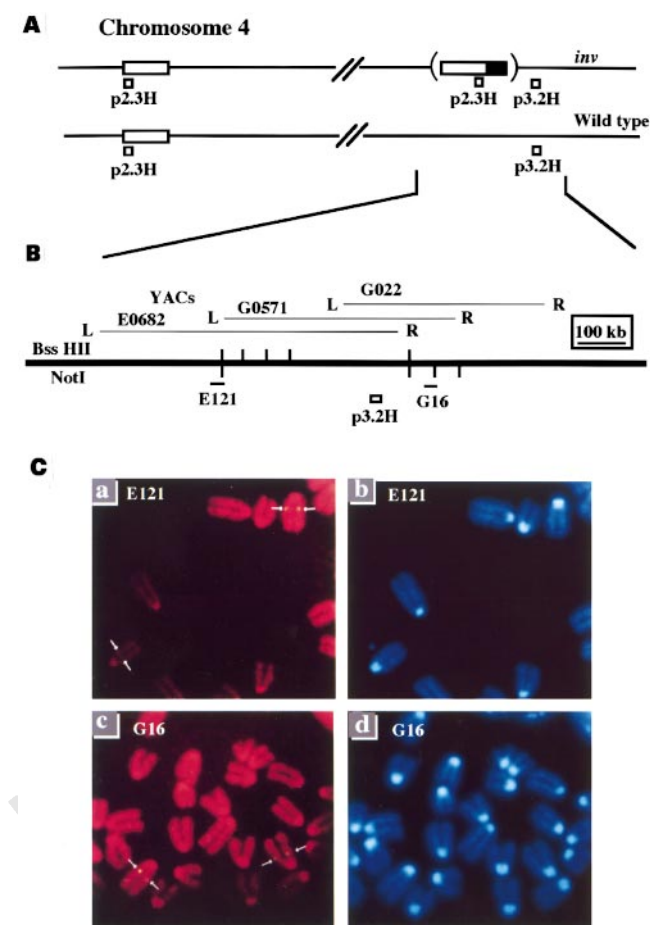


Figure 1 The *inv* transgenic integration site and FISH. **A**, Wild-type and *inv* mutant chromosomes 4. The open rectangle and black rectangle within the parentheses indicate the duplicated genomic region and the integrated tyrosinase minigene¹⁵, respectively. Probes p3.2H and p2.3H are shown as open squares under the lines. The duplicated regions on the *inv* genome separate 5–10% of the length of chromosome 4. **B**, YACs surrounding the transgenic integration site. The top three lines indicate three YACs (G022, G0571 and E0682). L, left arm of the YAC; R, right arm of the YAC. The bottom line shows the wild-type genome. Short vertical bars above and below the line indicate *Bss*III sites and *Not*I restriction sites, respectively. Short horizontal bars indicate cosmid clones E121 and G16. **C**, FISH. **a**, **c**, Mouse metaphase chromosomes from cultured fibroblasts of a heterozygous *inv* female mouse after FISH using cosmid E121 (**a**) or G16 (**c**) as a probe. **b**, **d**, The same chromosomes were stained with PI to identify chromosomes and their banding (**b**, **d**). All signals are located on chromosome 4 at band B (arrows).

of the *inv* chromosome¹⁵. Fluorescent *in situ* chromosomal hybridization (FISH) using E121 or G16 showed that both probes are located on chromosome 4 at band B of the wild-type and of the *inv* chromosome (Fig. 1C). The results do not support the presence of an intrachromosomal inversion in the *inv* genome, which had been proposed¹⁷ as a possible explanation for the *inv* phenotype.

We used genomic walking to determine the exact size of the deletion in the *inv* locus (Fig. 2a). A single-copy probe (G160 E1) is present in the *inv* genome, whereas probe H3-5 is missing, defining the extent of the genomic deletion between these two probes.

To identify transcribed regions in the genomic deletion, we screened a complementary DNA library derived from mouse embryos at embryonic day (E) 17.5 (Clontech) with cosmids E120 and EG5 (Fig. 2a) and identified one cDNA clone (clone 28). We screened a mouse cDNA library derived from E7.0 embryos

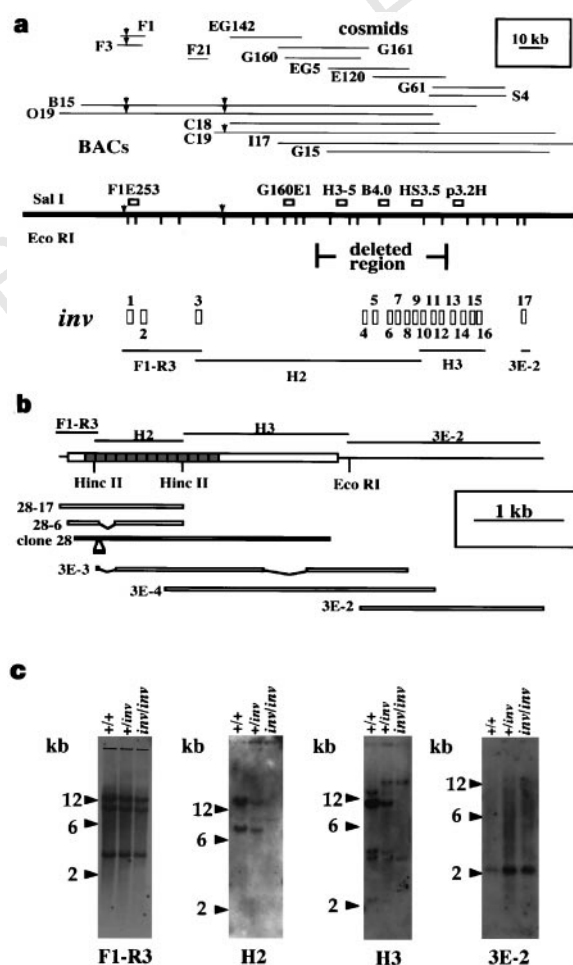


Figure 2 Genomic organization of the *inv* gene, cDNA clones and Southern hybridization. **a**, Phage, cosmid and BAC clones at the *inv* locus. The bold line indicates the wild-type genome. Lines above the bold line show BACs, cosmids and lambda clones (F1, F3, F21). Open rectangles above the bold line indicate probes for genomic walking. *Eco*RI and *Sal*I sites are shown as short vertical bars or arrows under or above the wild-type genome, respectively. The deleted region is also indicated below. Exons are shown as numbered boxes, and the lines beneath the exons indicate the regions of binding of the cDNA probes. **b**, Probes for Southern hybridizations, structure of the cDNA, and cDNA clones. Probes for Southern hybridizations are F1-R3, H2, H3 and 3E-2. In the cDNA structure, a boxed area indicates a coding region of the cDNA and a single line indicates an untranslated region. The shaded region indicates repeats of the Ank/Swi6 motif. The overlapping cDNA clones are 28-17, 28-6, clone 28, 3E-3, 3E-4 and 3E-2. Clone 28 contains extra sequence and 28-6 and 3E-3 lack sequences. **c**, Southern blot analyses of wild-type (+/+), heterozygous (+/inv) and homozygous (*inv/inv*) DNA with probes F1-R3, H2, H3 and 3E-2. Genomic DNAs were digested with *Eco*RI.

(Clontech) several times to obtain a full-length cDNA (Fig. 2b). The total size of the cDNA, based on the sequence, is ~5.6 kb. The size corresponds well to the 6.0-kb transcript size estimated from northern blot analysis (Fig. 3a, b).

We next analysed the genomic structure of this candidate gene. We screened a bacterial artificial chromosome (BAC) library (Research Genetics) with probe HS3.5 (ref. 15) and obtained six clones (B15, C18, C19, D5, G15, I17 and O19) (Fig. 1C). We also screened a C57Bl/6 genomic library with the 5' region of the cDNA and obtained three positive clones (F1, F3 and F21). These BACs and phages as well as cosmid clones were characterized by hybridization with genomic and cDNA probes, polymerase chain reaction (PCR) and pulsed-field gel analysis, and a physical map of the region surrounding the *inv* locus was made (Fig. 2a).

We used Southern blotting to hybridize parts of the cDNA to these BACs, cosmids and phages, and identified fragments in which exons of the candidate gene were located (Fig. 2a). Exon-intron boundaries were confirmed by sequencing.

To assay for missing exons in the *inv* mutant, we performed Southern hybridizations using four cDNA probes (Fig. 2), F1-R3, H2, H3 and 3E-2. Southern hybridizations with F1-R3 and 3E-2 showed that the *inv* homozygote retains the probed regions (Fig. 2c). Hybridizations using H2 and H3 showed that several bands are missing in the *inv* homozygote. Combined with the Southern hybridization pattern produced by using the flanking probe p3.2H (a wild-type band at 9.5 kb and the *inv* band at 20 kb, ref. 2), these hybridization results confirm that exons 4–12 of the candidate gene are missing in the *inv* genome (Fig. 2a).

Northern hybridization showed that the gene is expressed as early as E7 (Fig. 3a). Whole-mount *in situ* hybridization also showed that the candidate gene is expressed in presomite-stage embryos (Fig. 3c) before asymmetrical expression of *nodal* and *lef*. In the adult, the

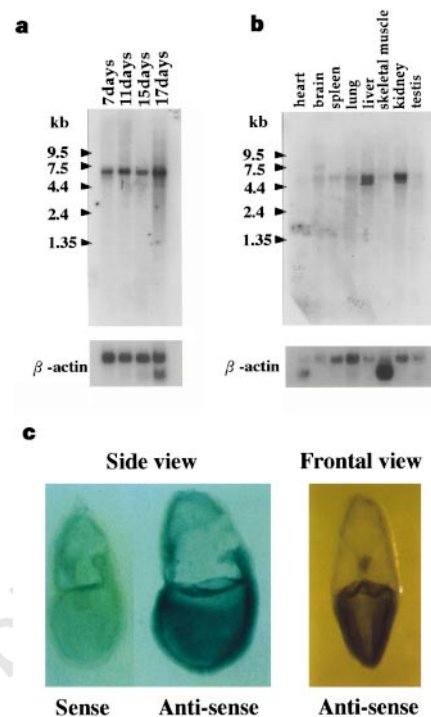


Figure 3 Northern blot analysis of embryonic (a) and adult (b) tissues of the mouse using the candidate *inv* cDNA. Clontech multiple tissue northern (MTN) blot membranes (a, CL7763-1; b, CL7762-1) were hybridized with a 32 P-labelled 3.0-kb *Eco*RI fragment of clone 28. Results of hybridization with a human β -actin probe are shown below. c, Whole-mount *in situ* hybridization to a presomite-stage mouse embryo.

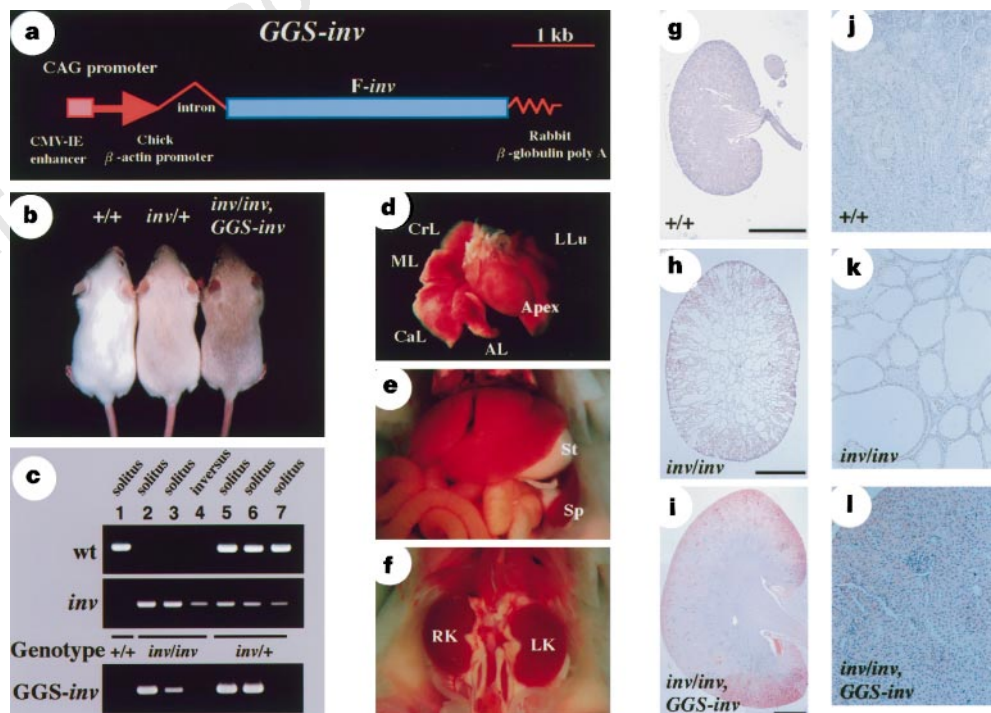


Figure 4 Transgenic rescue of *inv/inv* by introducing the candidate *inv* gene. a, The GGS-*inv* construct. b, Coat colour of 2-week-old mice from three different genotypes. c, Situs types and genotypes of seven pups. Genotypes were determined by PCR analysis of wild-type and *inv* alleles and of the transgene, GGS-*inv*. There are three *inv/inv* pups (2–4). Two of them (2, 3) contain a GGS-*inv* transgene and have the situs solitus phenotype, whereas the third (4) does not contain a GGS-*inv* transgene and has the situs inversus phenotype. d–f, Visceral organs of *inv/inv*, GGS-*inv* mice. All viscera in d, the thorax (lung, heart) and e, f,

the abdomen (liver, stomach, spleen and intestine (e) and kidney (f)) showed normal morphologies and positions. AL, accessory lobe; CaL, caudal lobe; CrL, cranial lobe; ML, medial lobe; LLu, left lung; Sp, spleen; St, stomach; RK, right kidney; LK, left kidney. g–i, Histological analysis of kidney (5 days post-natal) in a wild-type embryo (g, j), an *inv* homozygote (h, k), and an *inv* homozygote containing the GGS-*inv* transgene (5 weeks post-natal i, l). Scale bar indicates 1.0 mm (in g, h, k).


```

1  MNISEDVLST GSSLASQVHA AAVNGDKGAL QRLIVGNSAL RDKEDRFGRT PLMYCVLADR VDCADALLKA GADVNTDHS RRTALHLAAQ KGNRYFMKLL
      R1                                     R2
101 LTRRANWQK DLEEMTPLHL STRHRSPKCL ALLLKPMAPG EVDTQDKNQK TALHWSAYYN NPEHAKLLIK HDSNIGIPDV EGKIPHLWAA NHKDPSAVHT
      R3                                     R4                                     R5
201 VRCILDAAPT ESLLNWQDYE GRTPHLFAVA DGNLTVVDVL TSYESCNTS YDNLFRTPLH WAALLGHAQI VHLLERNKS GTIPSDSQGA TPLHYAAQSN
      R6                                     R7                                     R8
301 FAETVKVFLQ HPSVKDDSDL EGRTSFMWAA GKGNDVLRV MSLKSDID1 NMSDKYGGTA LHAAALSGHV STVKLLLDND AQVDATDVMK HTPLFRACEM
      R9                                     R10                                    R11
401 GHRDVIQTLI KCGARVDLVD QDGHSLJHWA ALGGNADVCQ ILIENKINPN VQDYAGRTPL QCAAYGGYIN CMAVLMENNA DPNIQDKEGR TALHWSSCNNG
      R12                                    R13                                    R14
501 YLDAIKLLLD FAAPFNQMEN NEERYTPLYD ALLGERHEVI QFMLEHGALS IAAIQDIAAF KIQAVYKGYK VRKAFRDRKN LLMKHEQLRK DAAAKKREEE
      R15
601 NKRKEAEQKQ QGLDTPPRS HCSSAPVLP CPPSPQNEGS QKDATPSKQP PASHTVQSPD PEHSRLPGRC PGRASQGDSS IDLQGTASRK PSETPIEHCR
701 GPSACVHPRS WEGGNSSKNQ GTSSVEKRRG ETNGKRRCE EGPSSARQPL CTGSGRPAEK GEDSSPAVAS ASQDHPKP NKRQDRAARR RGASQKRRTH
801 QLRDRCPAG SSRPGSAKE VACADQSSLH RHTPRSKVTQ DKLIGVSSG LPLSTEASRS GCKQLYEDIC ASPETGAHG PPPGQCMNH LLPVEQRLLI
901 IQRERSRKEL FRRKNKAAV IQRAWRSYQL RKHLRLLHL KQLGAREVLR CTQVCTALLL QVWRKELEK FPKSISVSRT SKSPSKGSSA TKYARHSVLR
1001 QIYGCSQEGK GHMFIKSSKA PAVLHLSVN SLQSIHLDNS GRSKKFSYNL QPSSQSKNKP KL

```

Figure 5 Deduced amino-acid sequence of the *inv* candidate gene. Underlining indicates 15 repeats of the Ank/Swi6 motif.

highest level of expression is seen in kidney and liver (Fig. 3b). Genomic analysis and expression patterns of the gene indicated that it is a good candidate for encoding the *inv* protein.

To determine whether the candidate gene can correct the phenotype seen in *inv* mutants, we generated transgenic mice carrying a minigene (GGs-*inv*) (Fig. 4a). As results of whole-mount *in situ* hybridization indicated that the gene may be expressed symmetrically and almost ubiquitously, we decided to use a CAG promoter, which consists of a cytomegalovirus immediate early (CMV-IE) enhancer and chick β -actin promoter¹⁸.

One GGS-*inv*-positive mouse (V15) was heterozygous for the *inv* mutation. This founder mouse (*inv*+/+, GGS-*inv*) was mated to *inv*/+ females to test whether the GGS-*inv* minigene could restore normal left/right asymmetry in *inv/inv* homozygotes. In the offspring resulting from these matings, three types of coat colour were produced, namely, albino, light brown and dark brown, which correspond to the +/+, *inv*/+ and *inv/inv* genotypes, respectively (Fig. 4b). Dark brown coat colour represents a gene-dosage effect of the tyrosinase minigene. These genotypes were confirmed by PCR (Fig. 4c).

We obtained 10 *inv* homozygotes among 53 offspring from 7 litters. Three of the homozygotes lacked the transgene and died after birth. The remaining seven *inv* homozygotes contained the GGS-*inv* transgene, and they all grew normally. We studied the visceral organs of three *inv/inv*, GGS-*inv* animals. In all cases, the laterality defects seen in the *inv* mutants were corrected (Fig. 4d–f). In the thorax, the right lung had more lobes than the left lung and the apex of the heart pointed to the left (Fig. 4d). Stomach and spleen (Fig. 4e) were located on the left side, and the right kidney was positioned anterior to the left one (Fig. 4f). Within a couple of days after birth, the kidneys of *inv* homozygous mice were filled with severely dilated collecting ducts (Fig. 4h, k) compared with normal kidneys (Fig. 4g, j). In contrast, the kidneys of the *inv/inv*, GGS-*inv* mouse developed normally (Fig. 4i, l). The transgenic rescue experiments show that the *inv* gene was identified correctly and that only one gene is responsible for the phenotype seen in the *inv* mutant.

The cDNA of the candidate gene has a long open reading frame which encodes 1,062 amino acids (Fig. 5). The most distinctive feature of this candidate protein is the presence of 15 successive repeats of the Ank/Swi6 motif, which consists of about 33 amino acids (Fig. 5). This motif has been found in vertebrate and invertebrate proteins, including ankyrins^{7,8,19}, developmental-regulatory proteins^{20,21} and cell-cycle-regulatory proteins^{22–24}. A BLAST

search showed that the candidate protein shares the highest homology with ankyrins^{7,8}. Like ankyrin, our candidate protein has a short amino-acid sequence before the Ank/Swi6 repeats. However, the candidate protein does not contain the spectrin-binding and regulatory domains found in ankyrins. We found no homology of the carboxy-terminal half of the candidate *inv* gene with other, previously described, genes. The candidate protein does not contain a signal peptide or transmembrane domain, indicating that it is probably an intracellular protein.

The genomic deletion in *inv* eliminates most of its coding sequences. Thus the *inv* mutant is probably a loss-of-function mutant. Another, less likely, possibility is that the remaining Ank/Swi6 repeat, either alone or fused with another protein, has a dominant-negative function.

Our transgenic rescue experiments indicate that localized expression of the *inv* protein may not be necessary, as the CAG promoter directs ubiquitous expression of the gene. As the *inv* protein is probably an intracellular protein, it may control the establishment of left/right asymmetry by affecting cell polarity rather than by its asymmetric distribution in embryos.

Injection of *Vg1* into the R3 cell in *Xenopus* embryos leads to a constant reversal of left/right asymmetry, and broadening the expression pattern of *Vg1* in aberrant cells on the right side of the embryos causes the *inv* phenotype^{25,26}. It is not known how this aberrant expression is induced in the absence of the *inv* protein.

Although the mechanisms by which the *inv* protein specifies left/right asymmetry and by which loss of the *inv* protein reverses left/right asymmetry still remain to be established, our identification of the gene involved will be useful in elucidating these mechanisms. □

Methods

Construction of cosmid libraries. Yeast DNA was partially digested with *Sau3AI* and was dephosphorylated and ligated with *Bam*HI–*Xba*I-digested sCos 1 vector (Stratagene). Ligated products were packaged using Gigapack II (Stratagene) and infected using NM554 or XL1BlueMR.

Suppressive hybridization for cDNA screening. Cosmid EG5 or E120 was digested with *NotI* and released from the vector. Inserts were labelled with [³²P]dCTP using the random priming method (Stratagene). ³²P-labelled probe was mixed with 50 μ g of mouse *Cot*I DNA (Gibco) and denatured by boiling. The mixture was incubated at 65 °C for 30 min and then hybridized to filters.

FISH analysis. Cosmid clones G16 and E121 were labelled by nick-translation with biotin-16-dUTP (Boehringer). Mouse chromosomes were prepared from early passages of fibroblasts from a heterozygous *inv* female mouse. Standard

FISH methods²⁷ were used with some modifications. Hybridized DNA probes were detected by fluorescence avidin DCS (Vector). We reduced background signals by competition with mouse *Cot1* sequences (added in hybridization solution). The slides were finally stained with 4',6'-diamino-2-phenylindole (DAPI) and propidium iodide (PI).

Whole-mount *in situ* hybridization. Embryos were collected from randomly bred ICR mice. To determine stages of development, we took the day on which a vaginal plug was detected to be E0.5. Digoxigenin-labelled (Boehringer) probe was generated with SP6 RNA polymerase (antisense) or T7 RNA polymerase (sense) from a 2.5-kb *EcoRI*–*HindIII* fragment of clone 28. Whole-mount *in situ* hybridizations were performed at high stringency²⁸.

Northern and Southern hybridization. Multiple-tissue northern blots (Clontech) were hybridized with clone 28, which was labelled by random-priming using the Prime it II kit (Stratagene) or an oligonucleotide labelling kit (Pharmacia) and [³²P]dCTP. Membranes were hybridized in express hybridization buffer (Clontech), washed with 2 × SSC (SSC is 0.15 M NaCl, 30 mM sodium citrate, pH 7.0), 0.1% SDS at 65 °C, and exposed overnight. For northern hybridization, the final wash was 2 × SSC, 0.1% SDS at 50 °C.

The GGS-*inv* construct and production and identification of transgenic mice. The GGS-*inv* transgene contains the full coding sequence of the *inv* candidate gene (*F-inv*) linked to the CAG promoter (Fig. 4a). A 6.0-kb fragment of GGS-*inv* was microinjected into embryos at the one-cell stage collected from *inv* heterozygous matings to generate transgenic mice, using standard techniques²⁹.

Genotypes were determined by Southern hybridization or by PCR analysis using the following primers: 11220 (5'-TGAGTTGTGTGGAGGAATCTCTT-3'), 11291 (5'-GTGATGTGCTGGAAGATGGAATTG-3'), 4633 (5'-GAC-TACCACAGCTTGCTGTATCAGAG-3'), B-28 (5'-TTTCTTCTAGAAGGAGATGGAC-3') and E-28 (5'-GCCGCTGCAGTTAATGGAGATAAA-3'). Primer pairs 4633 and 11220 were used to detect the *inv* mutant locus, 11291 and 11220 to detect the wild-type locus, and B-28 and E-28 to detect the GGS-*inv* transgene.

Tissue processing and histological analysis. Kidneys were fixed in Bouin's fixative, dehydrated and embedded in Paraplast. Serial sections (6-μm thickness) were stained with haematoxylin and eosin using standard procedures.

Received 21 July; accepted 17 August 1998.

- Yokoyama, T. *et al.* Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucleic Acids Res.* **18**, 7293–7298 (1990).
- Yokoyama, T. *et al.* Reversal of left–right asymmetry: a situs inversus mutation. *Science* **263**, 679–681 (1993).
- Meno, C. *et al.* Left–right asymmetric expression of the TGFβ-family member in mouse embryos. *Nature* **381**, 151–155 (1996).
- Meno, C. *et al.* Two closely-related left–right asymmetrically expressed genes, *lefty-1* and *lefty-2*: their distinct expression domains, chromosomal linkage and direct neuralizing activity in *Xenopus* embryos. *Genes Cells* **2**, 513–524 (1997).
- Collignon, J., Varlet, I. & Robertson, E. J. Relationship between asymmetric *nodal* expression and the direction of embryonic turning. *Nature* **381**, 155–158 (1996).
- Lowe, L. A. *et al.* Conserved left–right asymmetry of *nodal* expression and alterations in murine situs inversus. *Nature* **381**, 158–161 (1996).
- Lux, S. E., John, K. M. & Bennett, V. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**, 36–42 (1990).
- Bennett, V. Ankyrins, adapters between diverse plasma membrane proteins and the cytoplasm. *J. Biol. Chem.* **267**, 8703–8706 (1992).
- Levin, M., Johnson, R. L., Stern, C. D., Kuehn, M. & Tabin, C. A molecular pathway determining left–right asymmetry in chick embryogenesis. *Cell* **82**, 803–814 (1995).
- Isaac, A., Sargent, M. G. & Cooke, J. Control of vertebrate left–right asymmetry by a snail-related zinc finger gene. *Science* **275**, 1301–1304 (1997).
- Yost, H. J. Regulation of vertebrate left–right asymmetries by extracellular matrix. *Nature* **357**, 158–161 (1992).
- Hummel, K. P. & Chapman, D. B. Visceral inversion and associated anomalies in the mouse. *J. Hered.* **50**, 9–13 (1959).
- Layton, W. M. Random determination of a developmental process. Reversal of normal visceral asymmetry in the mouse. *J. Hered.* **67**, 336–338 (1976).
- Supp, D. M., Witte, D. P., Potter, S. S. & Brueckner, M. Mutation of an axonemal dynein affects left–right asymmetry in *inversus* situs mice. *Nature* **389**, 963–966 (1997).
- Yokoyama, T., Harrison, W., Elder, F. B. & Overbeek, P. A. in *Developmental Mechanisms of Heart Disease* (eds Clark, E. B. & Takao, A.) 513–520 (Futura, New York, 1995).
- Larin, Z., Monaco, A. P. & Lehrach, H. Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl Acad. Sci. USA* **88**, 4123–4127 (1991).
- Klar, A. J. S. A model for specification of the left/right axis in vertebrates. *Trends Genet.* **10**, 392–396 (1994).
- Niwa, H., Yamamura, K. & Miyazaki, J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193–200 (1991).
- Peter, L. L. & Lux, S. E. Ankyrins: structure and function in normal cells and hereditary spherocytes. *Semin. Hematol.* **30**, 85–118 (1993).
- Yochem, J., Weston, K. & Greenwald, I. The *Caenorhabditis elegans* *lin-12* gene encodes a transmem-

- brane protein with overall similarity to *Drosophila* *Notch*. *Nature* **335**, 547–550 (1988).
- Yochem, J. & Greenwald, I. *gfp-1* and *lin-12*, genes implicated in distinct cell–cell interactions in *C. elegans*, encode similar transmembrane proteins. *Cell* **58**, 887–898 (1989).
- Andrews, B. J. & Herskowitz, I. The yeast SWI4 protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle-dependent transcription. *Nature* **342**, 830–833 (1989).
- Aves, S. J., Durkacz, B. W., Carr, A. & Nurse, P. Cloning, sequencing and transcriptional control of the *Schizosaccharomyces pombe* *cdc10* 'start' gene. *EMBO J.* **4**, 457–463 (1985).
- Breeden, L. & Nasmyth, K. Similarity between cell-cycle control genes of budding yeast and fission yeast and the *Notch* gene of *Drosophila*. *Nature* **329**, 651–654 (1987).
- Hyatt, B. A., Lohr, J. A. & Yost, H. J. Initiation of vertebrate left–right axis formation by maternal *Vg1*. *Nature* **385**, 62–65 (1996).
- Hyatt, B. A. & Yost, H. J. The left–right coordinator; the role of *Vg1* in organizing left–right axis formation. *Cell* **93**, 37–46 (1998).
- Lawrence, J. B., Villnave, C. A. & Singer, R. H. Sensitive, high resolution chromatin and chromosome mapping *in situ*: presence and orientation of two closely integrated copies of EBV in a lymphoma cell line. *Cell* **52**, 51–61 (1988).
- Wilkinson, D. G. in *In Situ Hybridization: A Practical Approach* (ed. Wilkinson, D. G.) 75–84 (IRL, Oxford, 1992).
- Hogan, B., Costantini, F. & Lacy, L. in *Manipulating the Mouse Embryo: A Laboratory Manual* 153–203 (Cold Spring Harb. Lab., Cold Spring Harbor, 1986).

Acknowledgements. We thank A. Takao for encouragement and support; R. Kucherlapati, S. Somlo and S. Aizawa for suggestions; H. Lehrach for YAC clones; M. Ishibashi and M. Yokoyama for mouse husbandry and embryo preservation; S. Ohishi and K. Mochida for mouse genotyping and histological sections; the staff of the TWMC for supporting our projects; and J. Miyazaki for providing the pCAGGS vector.

Correspondence and requests for materials should be addressed to T.Y. (e-mail: tyoko@research.twmc.ac.jp). The *inv* gene sequence has been deposited with Genebank (accession number AF034860).

Zebrafish organizer development and germ-layer formation require nodal-related signals

Benjamin Feldman*, Michael A. Gates*, Elizabeth S. Egan*, Scott T. Dougan*, Gabriela Rennebeck*, Howard I. Sirotkin, Alexander F. Schier & William S. Talbot

Developmental Genetics Program, Skirball Institute of Biomolecular Medicine and Department of Cell Biology, New York University Medical Center, 540 First Avenue, New York, New York 10016, USA

* These authors contributed equally to this work

The vertebrate body plan is established during gastrulation, when cells move inwards to form the mesodermal and endodermal germ layers. Signals from a region of dorsal mesoderm, which is termed the organizer, pattern the body axis by specifying the fates of neighbouring cells^{1,2}. The organizer is itself induced by earlier signals¹. Although members of the transforming growth factor-β (TGF-β) and Wnt families have been implicated in the formation of the organizer, no endogenous signalling molecule is known to be required for this process¹. Here we report that the zebrafish *squint* (*sqt*)³ and *cyclops* (*cyc*)⁴ genes have essential, although partly redundant, functions in organizer development and also in the formation of mesoderm and endoderm. We show that the *sqt* gene encodes a member of the TGF-β superfamily that is related to mouse *nodal*. *cyc* encodes another nodal-related protein^{5,6}, which is consistent with our genetic evidence that *sqt* and *cyc* have overlapping functions. The *sqt* gene is expressed in a dorsal region of the blastula that includes the extraembryonic yolk syncytial layer (YSL). The YSL has been implicated as a source of signals that induce organizer development and mesendoderm formation^{2,7}. Misexpression of *sqt* RNA within the embryo or specifically in the YSL induces expanded or ectopic dorsal mesoderm. These results establish an essential role for nodal-related signals in organizer development and mesendoderm formation.

We identified a mutation in *sqt*³ as a spontaneous recessive enhancer of the *cyc*⁴ mutant phenotype (see Methods). *sqt;cyc* double mutants lack mesendodermal derivatives, including notochord,

somites, heart, pronephros, blood and gut, in the head and trunk (Fig. 1, and data not shown). In comparison to *sqt;cyc* double mutants, the phenotypes of *sqt* and *cyc* single mutants are much less severe. Each single mutant displays cyclopia and defects in prechordal plate and ventral nervous system^{3,4,8}. The phenotype of the double mutant thus reveals partly overlapping functions of the *sqt* and *cyc* genes in the development of mesoderm and endoderm.

To characterize the functions of the *sqt* and *cyc* genes in the formation and patterning of mesoderm and endoderm, we examined *sqt;cyc* mutants at blastula and gastrula stages. An early step in dorsal mesoderm development fails in *sqt;cyc* mutants, as indicated by a lack of the dorsal mesodermal marker *gooseoid* (*gsc*)⁹ (Fig. 1h). Furthermore, dorsal expression of the pan-mesodermal marker *ntl/brachyury*¹⁰ is lacking in *sqt;cyc* mutants (Fig. 1j), whereas ventral expression is evident from late blastula until the end of gastrulation,

when transcripts are present in the tail bud (Fig. 1l). Expression of *gsc* is also reduced in *sqt* single mutants at blastula stages, but in contrast to *sqt;cyc* double mutants, *gsc* expression increases as gastrulation progresses (data not shown). These early disruptions of marker gene expression are followed by morphologically apparent defects at later stages: the zebrafish organizer region², the dorsal thickening called the embryonic shield, is not evident at the onset of gastrulation in *sqt* single mutants (Fig. 1d) or *sqt;cyc* double mutants (not shown). This analysis shows that the *sqt* and *cyc* genes have an early zygotic function in dorsal mesoderm induction.

In addition to the absence of dorsal mesoderm in the blastula, we observed extensive defects in mesoderm and endoderm formation in *sqt;cyc* mutants during gastrulation. Markers of axial, paraxial, intermediate and ventral mesoderm are not expressed (Fig. 1k–r,

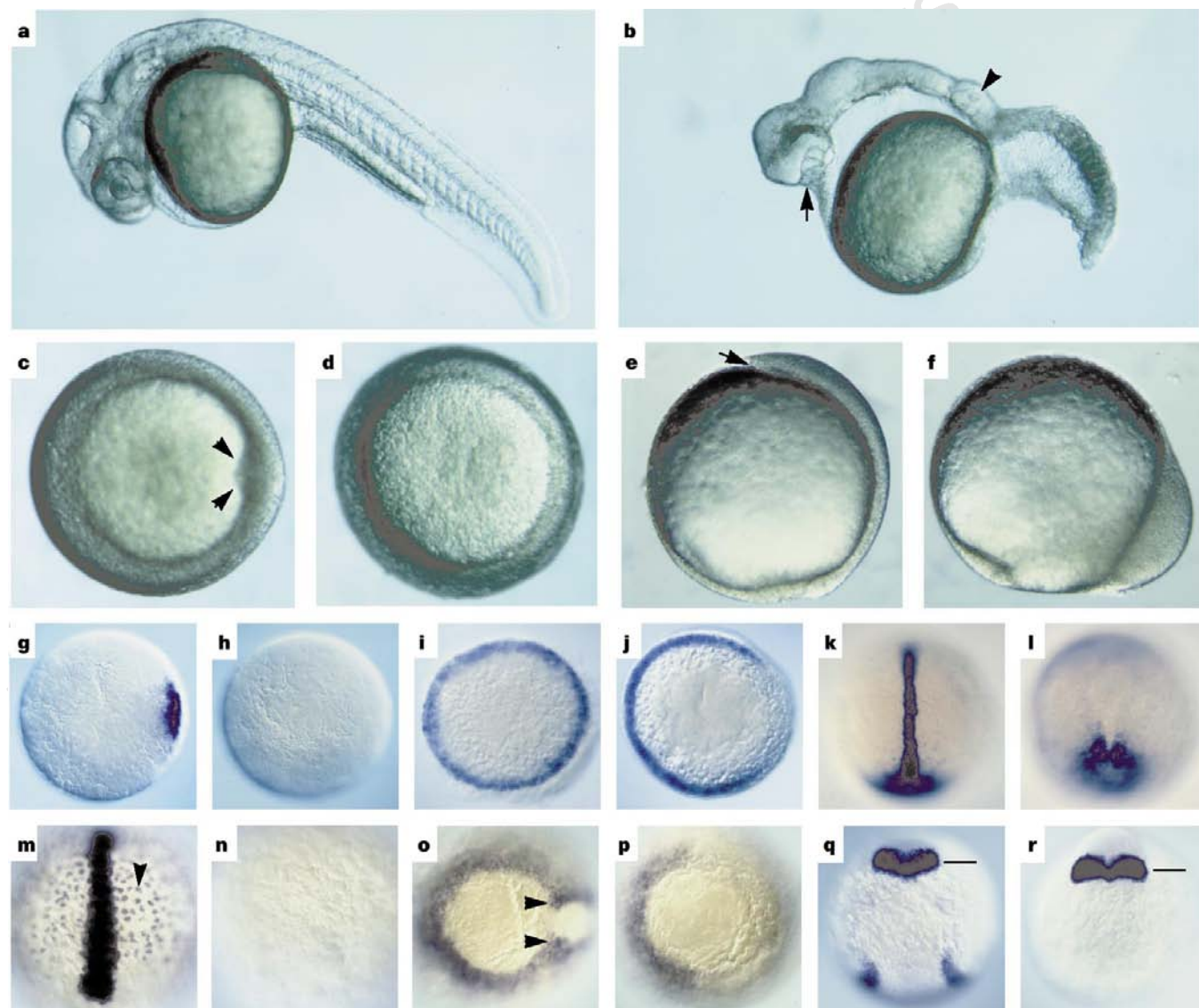


Figure 1 Disruption of mesoderm and endoderm in *sqt;cyc* double mutants. Wild-type (a, c, e, g, i, k, m, o, q), *sqt;cyc* double mutant (b, f, h, j, l, n, p, r), and *sqt* single mutant (d) embryos are shown. a, b, Phenotypes at 1 day after fertilization (side view, anterior to the left). Arrow in b marks the lens of the eye, and the arrowhead indicates the otic vesicle. c, d, Phenotypes of shield-stage embryos (6 h); animal-pole views. Arrows in c mark the embryonic shield in the wild-type embryo. The shield is not seen in *sqt* (d) or *sqt;cyc* (not shown) mutants at this stage. e, f, Phenotypes at the late gastrula stage (95% epiboly); lateral views, dorsal to the right. Arrowhead in e points to the anterior limit of the hypoblast layer

in the wild-type embryo. g–r, Expression of g, h, *gooseoid* (dorsal mesoderm)^{8,9}, i–l, *ntl/brachyury* (presumptive mesoderm, notochord, tail bud)¹⁰; m, n, *axial/HNF-3β* (axial mesoderm, endoderm (arrowhead))^{11,12}; o, p, *snail-1* (presumptive mesoderm, adaxial cells (arrowhead))¹³; and q, r, *pax2.1* (intermediate mesoderm, mid-hindbrain boundary (lines))¹⁵. Panels g–j show animal-pole views at the 30% epiboly (4.7 h) stage. Panels k–n, q and r show dorsal views at bud (10 h; k, l, q and r) and 80% epiboly (8.5 h; m and n) stages. Panels o and p show vegetal-pole views, dorsal to the right, at the 80% epiboly (8.5 h) stage.

and data not shown). Moreover, endodermal expression of the *axial/HNF-3 β* gene^{11,12} is absent (Fig. 1n). Exceptions to the general failure of mesoderm formation occur in the tail, where variable expression of α -tropomyosin¹³ indicates that muscle differentiation can occur (not shown). In addition to the disruptions in mesodermal and endodermal gene expression, the morphogenesis of these germ layers is abnormal. The failure of marginal cells to involute, or move inwards, in *sqt;cyc* double mutants results in the absence of the hypoblast, the embryonic layer comprising the mesoderm and endoderm¹⁴. The movements of epiboly and convergence, which drive cells vegetally and dorsally¹⁴, respectively, do occur, leading to the accumulation of cells in the dorsal-vegetal region of the late gastrula (Fig. 1f).

In contrast to mesendodermal derivatives, ectodermal structures form in *sqt;cyc* mutants. Considering the absence of mesodermal and endodermal derivatives, the degree of anteroposterior neural patterning is particularly striking: polarity is indicated both by morphological features such as the single median eye at the anterior end of the neural tube (Fig. 1b) and by molecular markers such as *pax2.1* (ref. 15), which is expressed at the midbrain–hindbrain boundary in wild type and *sqt;cyc* mutants (Fig. 1r). This indicates that vertical signals from involuted dorsal mesoderm are not required for anteroposterior pattern in the neuraxis of *sqt;cyc* mutants; it is possible that planar signals from two other organizing centres in zebrafish (reviewed in ref. 2) induce this pattern.

To initiate a molecular analysis, we genetically mapped *sqt*. We

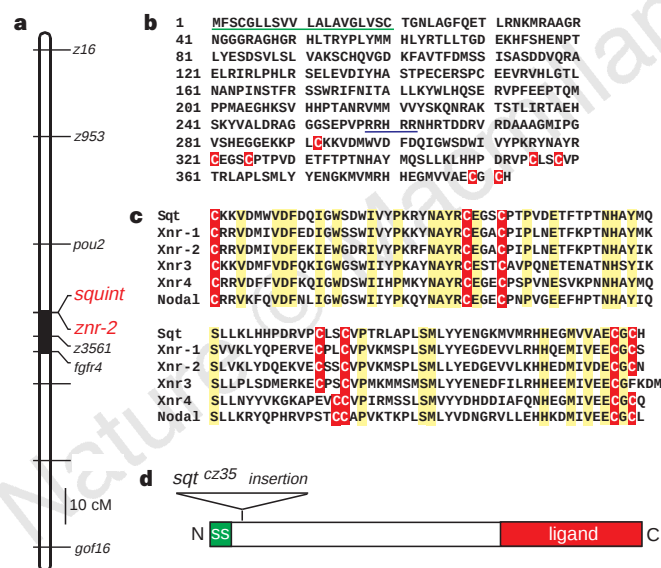


Figure 2 The *sqt* gene encodes a nodal-related signal. **a**, Genetic map of LG 21 showing the location of the *sqt* locus, the *znr-2* gene and nearby markers. Polymorphisms in the *znr-2* and *fgfr-4* genes²⁶ and the marker z3561 (ref. 27) were scored in *sqt* mapping crosses; the positions of other loci shown and the centromere (filled) are inferred from other maps^{26,27} constructed with wild-type mapping panels. **b**, Predicted amino-acid sequence of the Sqt protein. The predicted N-terminal signal sequence (green) and basic cleavage sites (purple) are underlined. The seven conserved Cys residues that are characteristic of the mature ligand domain of TGF- β superfamily members are highlighted in red. **c**, Comparison of Sqt and nodal-related proteins^{17,18,28,29} from frog and mouse, including Xnr-1, Xnr-2, Xnr-3, Xnr-4 and nodal. Alignment begins at the first conserved Cys residue (red) of the mature ligand domain. Amino acids that are identical in the six proteins are highlighted (yellow). The Sqt, Xnr-1, Xnr-2 and Xnr-3 proteins share an unusual 'CXXC' motif that separates the fourth and fifth Cys residues, whereas nodal and Xnr-4 have the more prevalent 'CC' motif. **d**, Schematic representation of the Sqt protein, showing positions of the signal sequence, the mature ligand domain, and the insertion in the *sqt*^{cz35} mutation. This lesion disrupts the ORF after codon 30 and truncates the polypeptide 83 codons thereafter. The ~1.9 kb insert is not drawn to scale.

Table 1 *sqt* RNA induces *gsc* expression in wild type and *sqt* mutants

RNA injected	goosecoid expression pattern			
	Wild type	<i>sqt</i> mutant	Expanded	Ectopic
<i>lacZ</i>	32 (68%)	15 (32%)	0	0
<i>sqt</i> (5 pg)	16 (21%)	9 (12%)	41 (53%)	12 (15%)
<i>sqt</i> (10 pg)	0	0	75 (95%)	4 (5%)

found that the *sqt*^{cz35} mutation maps to linkage group 21 (LG 21) near zebrafish *nodal-related-2* (*znr-2*, C. Erter and C. V. E. Wright, GenBank accession number AF003699; also called *ndr1* (ref. 16), making this gene a candidate for the *sqt* locus. By using a probe derived from a partial *znr-2* sequence, we identified a full-length complementary DNA clone. Conceptual translation of the 1,176 base pair (bp) open reading frame (ORF) revealed a protein of 392 amino acids with the characteristics of the nodal-related^{17,18} subgroup of the TGF- β superfamily. These features include a predicted amino-terminal signal sequence, a conserved basic cleavage site, and a putative mature signalling domain with seven conserved cysteines in an arrangement that is characteristic of some nodal-related proteins¹⁸ in lower vertebrates (Fig. 2b–d). A search for molecular lesions in the *sqt*^{cz35} mutant allele of *znr-2* identified a 1.9 kilobase (kb) insertion that disrupts the gene after codon 30 of the ORF and leads to translational termination 83 codons thereafter (Fig. 2d). By assaying this insertion in *sqt* crosses, we confirmed that *znr-2* is tightly linked to *sqt* (0 recombinants among 111 meioses). Micro-injection of synthetic *znr-2* RNA restores *gsc* expression in *sqt* mutants (Table 1 and Fig. 3), providing additional evidence that *znr-2* is the *sqt* gene, and we refer to it as *sqt* hereafter. The *cyc* gene also encodes a nodal-related protein^{5,6}, and this, together with the genetic evidence for partly redundant functions, indicates that the Sqt and Cyc proteins have similar or identical activities.

The expression pattern of *sqt* supports an early function in the induction of dorsal mesoderm and germ-layer formation. After maternal expression (data not shown), the *sqt* gene is expressed in a dorsal sector of the midblastula margin, where the organizer will later form (Fig. 4a–d). Dorsal expression is evident as early as the 1,000-cell stage (3 h; Fig. 4a), before the appearance of other dorsally expressed genes, including *gsc*, *ntl/brachyury* and *axial/HNF-3 β* (refs 8–12). Transcripts are detectable in embryonic cells of the blastoderm and also in the extraembryonic YSL (Fig. 4d), which has been implicated as a source of signals that induce organizer development and mesendoderm formation^{2,7}. By the late blastula stage, *sqt* mRNA accumulates in cells around the circumference of the blastoderm margin, the location of mesendodermal precursors¹⁴ (Fig. 4e, f). In the early gastrula, *sqt* expression becomes restricted to a group of superficial cells at the dorsal blastoderm margin (Fig. 4g, h). By the end of gastrulation, *sqt* expression is not detectable.

To determine whether *sqt* can induce dorsal mesoderm, we examined *gsc* expression in embryos injected with synthetic *sqt* mRNA. Embryos injected in one blastomere at the 8–32-cell stage displayed expanded or, less frequently, ectopic *gsc* expression (Fig. 3 and Table 1). Injection of *sqt* RNA into the YSL at the 1,000-cell stage also induced expanded and ectopic *gsc* expression (Fig. 3), indicating that extraembryonic expression of *sqt* can induce dorsal mesoderm formation in the overlying blastoderm.

Our results establish an essential role of zebrafish nodal-related signals in mesoderm and endoderm formation during gastrulation. Previous work has suggested roles in mesoderm induction for TGF- β family members, including activin, Vg1, nodal, Xnr-1 and Xnr-2 (refs 1 and 17–20 and references therein). Mutational analysis in mouse showed that activin is not required for mesoderm induction²¹, but mouse *nodal* mutants lack mesoderm, indicating an essential role in mesoderm development²⁰. In addition to a requirement for mesoderm development, we find an essential

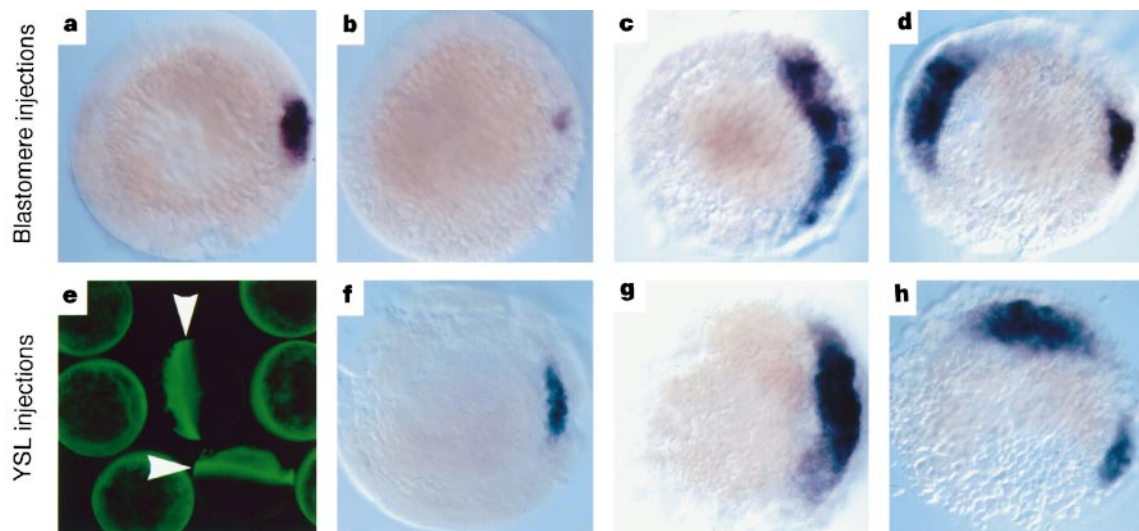


Figure 3 Microinjection of *sqt* RNA induces *gsc* expression in wild type and *sqt* mutants. **a-d**, Expression of *gsc* at 40% epiboly (5 h) in embryos from *sqt*^{c235/+} intercrosses that were injected in one blastomere at the 8–32-cell stage. *sqt* mutants (**b**) injected with *lacZ* are readily distinguished by their reduced *gsc* expression relative to wild-type embryos (**a**) injected with *lacZ*. Microinjection of *sqt* RNA induces *gsc* expression in wild-type and *sqt* mutant embryos, and examples of embryos with expanded (**c**) and ectopic (**d**) *gsc* expression are

shown. The frequencies of these phenotypic classes are shown in Table 1. **e-g**, YSL injections into wild-type embryos. **e**, Co-injected fluorescein-dextran lineage tracer was monitored at dome stage to verify localization to the YSL. Portions of eight YSL-injected embryos are shown. Two (arrowheads) are lateral views, and the others are animal-pole views. **f-h**, Expression of *gsc* at shield stage (6 h) in embryos injected in the YSL at the 1,000-cell stage with *lacZ* RNA (**f**) or *sqt* RNA (**g, h**).

function in endoderm formation, pointing toward a general role for nodal-related factors in signalling interactions that specify the fates of marginal blastomeres, or perhaps in inducing the involution movements of these cells. Exceptions to the general failure of mesoderm formation in *sqt;cyc* mutants occur in the developing tail, where some mesodermal cells undergo morphogenetic movements that are distinct from involution²².

Our studies also establish an essential role for nodal-related signals in the induction of dorsal mesoderm and in organizer development. Evidence indicates that the organizer is induced by a dorsalizing signal that acts downstream of maternally supplied β -

catenin^{1,2}. The zebrafish yolk cell can induce dorsal mesoderm⁷, and this activity is correlated with nuclear accumulation of β -catenin in the dorsal YSL and marginal blastomeres²³. Based on the absence of dorsal mesoderm in *sqt;cyc* mutants, the expression of the *sqt* gene in the dorsal YSL and marginal blastomeres, and the ability of *sqt* RNA to induce ectopic *gsc* expression, we propose that nodal-related signals function zygotically to relay patterning information that is established initially by maternal determinants such as β -catenin. Neural induction and dorsalward convergence take place in *sqt;cyc* mutants, indicating that other dorsalizing signals act in parallel to *sqt* and *cyc*. □

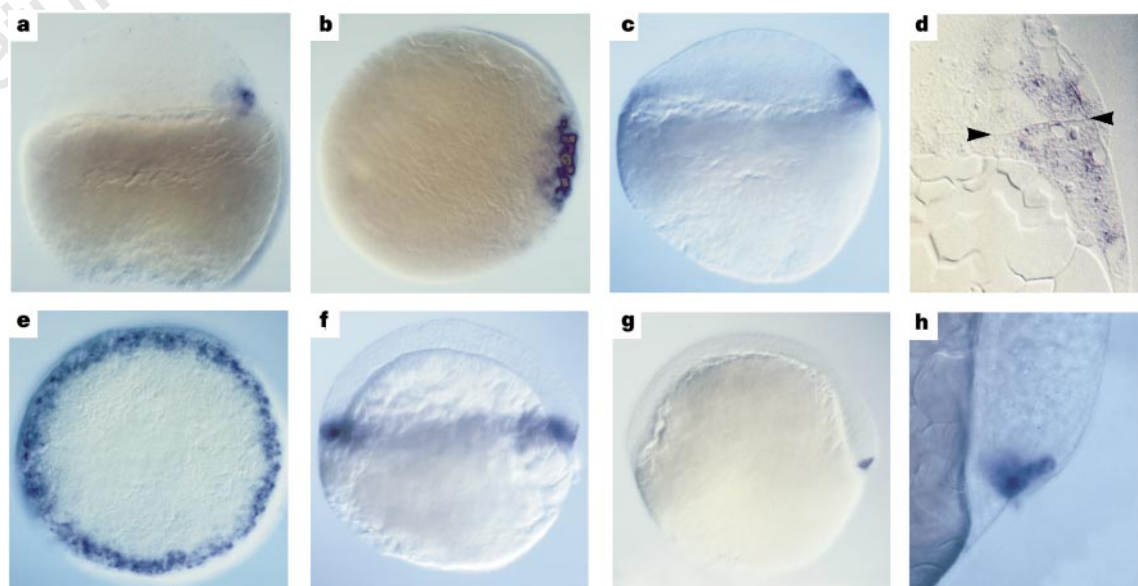


Figure 4 Expression of *sqt* mRNA in embryos at blastula and gastrula stages. *In situ* hybridization showing *sqt* expression at the following stages: **a**, 1,000 cell (3 h); **b**, high (3.3 h); **c, d**, sphere (4 h); **e**, 30% epiboly (4.7 h); **f**, germ ring (5.7 h); and **g, h**, shield (6 h) stages. **d**, Section showing *sqt* expression in the dorsal YSL and

dorsal blastoderm; arrows mark the YSL–blastoderm boundary. **h**, The shield region, showing superficial expression. **a, c, f–h**, Lateral views; **b, e**, Animal-pole views. Dorsal is to the right in **a–d, g** and **h**.

Methods

Fish stocks and phenotype characterization. We identified a mutation in *sqt*³ as a spontaneous recessive enhancer of the *cyc* mutant phenotype in a DAR × *cyc*^{m294} cross²⁴. This mutation is identical to *sqt*^{cz35} (which was formerly known as *sqt*^{z1} (ref. 3)), based on genetic mapping and the observation that both mutations contain the same insertion in the *sqt* gene as determined by a polymerase chain reaction (PCR) test with insertion-specific primers. The *sqt*^{cz35} mutation was identified initially as a spontaneous mutation in a DAR F1 cross (B. Paw and L. Zon, personal communication), which is consistent with the evidence that our mutation is a re-isolate of *sqt*^{cz35}. *In situ* hybridization was done as described previously^{12,13}, using staged¹⁴ embryos derived from natural matings of *sqt*^{cz35/+}; *cyc*^{m294/+} or *sqt*^{cz35/+} fish. The *cyc*^{m294} (ref. 25) allele behaves as amorph in genetic tests²⁴, and molecular analysis predicts that *sqt*^{cz35} is a null allele (Fig. 2). As expected for a phenotype caused by inheritance of two mendelian recessive mutations, about one-sixteenth (81 among 1,269 embryos) of the progeny in *sqt*^{cz35/+}; *cyc*^{m294/+} intercrosses are *sqt*; *cyc* mutants. For experiments in which mutants were not identified by their morphological phenotypes (for example, when analysis was done at early stages), the phenotypic class representing *sqt*; *cyc* mutants was inferred from its frequency.

Genetic mapping. Using standard methods^{24,26}, we mapped the *sqt* locus to the vicinity of the simple-sequence length polymorphism z3561 (five recombinants among 129 meioses). We mapped *znr-2* to LG 21 in a haploid mapping panel²⁴ by scoring a single-strand conformation polymorphism with primers (5'-GTTTATCCAAAGCGTTACAACGC-3' and 5'-TGTCTCATGACCATCTT GCCATT-3') that were designed from a partial *znr-2* sequence (C. Erter and C. V. E. Wright, GenBank accession number AF003699). To examine the linkage of *znr-2* in *sqt*^{cz35} mapping crosses, reverse primers detecting the presence (5'-ATATAAATCAGTACAACCGCCCG-3') or absence (5'-GCCAGCTGCTCGC ATTTTATTCC-3') of the insertion were used with a forward primer (5'-GAGCTTTATTTCATAACTGCGTG-3') present in wild-type and *sqt*^{cz35} alleles.

Sequence analysis and molecular characterization of *sqt* mutation. A full-length *znr-2* cDNA clone was obtained by screening a library from blastula and gastrula stages (T. Lepage and D. Kimelman, personal communication) with a probe synthesized with the primers described above. Both strands of the cDNA insert were sequenced with an ABI 377 automated sequencer. To search for molecular lesions in the *sqt*^{cz35} mutation, we used primers corresponding to the *znr-2* gene to amplify fragments from genomic DNA of wild-type and mutant siblings.

Microinjection of *squint* RNA. To assay for induction of dorsal mesoderm in wild type and *sqt* mutants, single blastomeres of embryos at the 8–32-cell stage from *sqt*^{cz35/+} intercrosses were co-injected with 5 pg or 10 pg of capped, synthetic *sqt* mRNA and 25 pg of capped, synthetic *lacZ* mRNA. Control embryos were injected with 25 pg of *lacZ* mRNA alone. Expression of *gsc* was examined at the 40% epiboly stage (5 h), and embryos were classified by *gsc* expression pattern (Fig. 3a–d). To assay the ability of *sqt* to act in the YSL, 25 pg of either *sqt* or *lacZ* mRNA was co-injected with 5% fluorescein–dextran into the YSL of wild-type embryos at the 1,000-cell stage (3 h). Injected embryos were examined with fluorescein optics to verify YSL localization of the lineage tracer. Embryos showing the YSL pattern illustrated in Fig. 3 (225 of 234 injected) were fixed at shield stage (6 h) and processed to examine *gsc* expression.

Received 23 April; accepted 28 July 1998.

- Harland, R. & Gerhart, J. Formation and function of Spemann's organizer. *Annu. Rev. Cell Dev. Biol.* **13**, 611–667 (1997).
- Schier, A. F. & Talbot, W. S. The zebrafish organizer. *Curr. Opin. Genet. Dev.* **8**, 464–471 (1998).
- Heisenberg, C. P. & Nüsslein-Volhard, C. The function of *silberblick* in the positioning of the eye anlage in the zebrafish embryo. *Dev. Biol.* **184**, 85–94 (1997).
- Hatta, K., Kimmel, C. B., Ho, R. K. & Walker, C. The *cyclops* mutation blocks specification of the floor plate of the zebrafish central nervous system. *Nature* **350**, 339–341 (1991).
- Sampath, K. et al. Induction of the zebrafish ventral brain and floorplate requires *cyclops*/nodal signalling. *Nature* (in the press).
- Rebagliati, M. R., Toyama, R., Haffter, P. & Dawid, I. B. *Cyclops* encodes a nodal-related factor involved in midline signalling. *Proc. Natl Acad. Sci. USA* **95**, 9932–9937 (1998).
- Mizuno, T., Yamaha, E., Wakahara, M., Kuroiwa, A. & Takeda, H. Mesoderm induction in zebrafish. *Nature* **383**, 131–132 (1996).
- Thisse, C., Thisse, B., Halpern, M. E. & Postlethwait, J. H. *gooseoid* expression in neuroectoderm and mesoderm is disrupted in zebrafish *cyclops* gastrulas. *Dev. Biol.* **164**, 420–429 (1994).
- Stachel, S. E., Grunwald, D. J. & Myers, P. Z. Lithium perturbation and *gooseoid* expression identify a dorsal specification pathway in the pregastrula zebrafish. *Development* **117**, 1261–1274 (1993).
- Schulte-Merker, S., Ho, R. K., Herrmann, B. G. & Nüsslein-Volhard, C. The protein product of the

zebrafish homologue of the mouse *T* gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* **116**, 1021–1032 (1992).

- Strähle, U., Blader, P., Henrique, D. & Ingham, P. W. *Axial*, a zebrafish gene expressed along the developing body axis, shows altered expression in *cyclops* mutant embryos. *Genes Dev.* **7**, 1436–1446 (1993).
- Schier, A. F., Neuhauss, S. C. F., Helde, K. A., Talbot, W. S. & Driever, W. The *one-eyed pinhead* gene functions in mesoderm and endoderm formation in zebrafish and interacts with *no tail*. *Development* **124**, 327–342 (1997).
- Thisse, C., Thisse, B., Schilling, T. F. & Postlethwait, J. H. Structure of the zebrafish *snail1* gene and its expression in wild-type, *spadetail* and *no tail* mutant embryos. *Development* **119**, 1203–1215 (1993).
- Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
- Krauss, S., Johansen, T., Korzh, V. & Fjose, A. Expression of the zebrafish paired box gene *pax[zf-b]* during early neurogenesis. *Development* **113**, 1193–1206 (1991).
- Rebagliati, M. R., Toyama, R., Fricke, C., Haffter, P. & Dawid, I. B. Zebrafish nodal-related genes are implicated in axial patterning and establishing left-right asymmetry. *Dev. Biol.* **199**, 261–272 (1998).
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. & Kuehn, M. R. Nodal is a novel TGF-β-like gene expressed in the mouse node during gastrulation. *Nature* **361**, 543–547 (1993).
- Jones, C. M., Kuehn, M. R., Hogan, B. M. L., Smith, J. C. & Wright, C. V. E. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651–3662 (1995).
- Toyama, R., O'Connell, M. L., Wright, C. V. E., Kuehn, M. R. & Dawid, I. B. Nodal induces ectopic *gooseoid* and *lim1* expression and axis duplication in zebrafish. *Development* **121**, 383–391 (1995).
- Conlon, F. L. et al. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919–1928 (1994).
- Matzuk, M. M. et al. Functional analysis of activins during mammalian development. *Nature* **374**, 354–356 (1995).
- Kanki, J. P. & Ho, R. K. The development of the posterior body in zebrafish. *Development* **124**, 881–893 (1997).
- Schneider, S., Steinbeisser, H., Warga, R. M. & Hausen, P. β-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mech. Dev.* **57**, 191–198 (1996).
- Talbot, W. S. et al. Genetic analysis of chromosomal rearrangements in the *cyclops* region of the zebrafish genome. *Genetics* **148**, 373–380 (1998).
- Schier, A. F. et al. Mutations affecting the development of the embryonic zebrafish brain. *Development* **123**, 165–178 (1996).
- Postlethwait, J. H. et al. Vertebrate genome evolution and the zebrafish gene map. *Nature Genet.* **18**, 345–349 (1998).
- Knapik, E. W. et al. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genet.* **18**, 338–343 (1998).
- Smith, W. C., McKendry, R., Ribisi, S. & Harland, R. M. A nodal-related gene defines a physical and functional domain within the Spemann organizer. *Cell* **82**, 37–46 (1995).
- Joseph, E. M. & Melton, D. A. *Xnr4*: a *Xenopus nodal*-related gene expressed in the Spemann organizer. *Dev. Biol.* **184**, 367–372 (1997).

Acknowledgements. We thank C. Erter, C. Wright, M. Rebagliati and I. Dawid for sharing and allowing us to cite their unpublished data; members of the Talbot and Schier laboratories for discussions; T. Lepage, D. Kimelman and C.-P. Heisenberg for reagents and fish stocks; S. McManus for fish care; and A. Ruiz i Altaba and G. Fishell for comments on the manuscript. We acknowledge postdoctoral fellowship support from the NIH (B.F. and H.L.S.) and ACS (S.T.D.). This work was supported by grants from the NIH (W.S.T. and A.F.S.) and an NYU Whitehead Fellowship (W.S.T.).

Correspondence and requests for materials should be addressed to W.S.T. (e-mail: talbot@saturn.med.nyu.edu). The *squint* cDNA sequence has been deposited in the GenBank database under accession no. AF056327.

Induction of the zebrafish ventral brain and floorplate requires *cyclops*/nodal signalling

Karuna Sampath*†‡, Amy L. Rubinstein‡, Abby M. S. Cheng*‡, Jennifer O. Liang‡, Kimberly Fekany§, Lilianna Solnica-Krezel§, Vladimir Korzh†, Marnie E. Halpern‡ & Christopher V. E. Wright*

* Department of Cell Biology, Vanderbilt University School of Medicine, and § Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37232, USA

† Institute of Molecular Agrobiolgy, National University of Singapore, 1 Research Link, 117604, Singapore

‡ Department of Embryology, Carnegie Institution of Washington, Baltimore, Maryland 21210, USA

‡ These authors contributed equally to this work

Zebrafish *cyclops* (*cyc*) mutations cause deficiencies in the dorsal mesendoderm^{1,2} and ventral neural tube^{3,4}, leading to neural defects and cyclopia^{5,6}. Here we report that *cyc* encodes a transforming growth factor-β (TGF-β)-related intercellular signalling molecule that is similar to mouse *nodal*⁷. *cyc* is expressed in dorsal mesendoderm at gastrulation and in the prechordal plate until

early somitogenesis. Expression reappears transiently in the left lateral-plate mesoderm, and in an unprecedented asymmetric pattern in the left forebrain. Injection of *cyc* RNA non-autonomously restores *sonic hedgehog*-expressing cells of the ventral brain and floorplate that are absent in *cyc* mutants, whereas inducing activities are abolished by *cyc*^{m294}, a mutation of a conserved cysteine in the mature ligand. Our results indicate that *cyc* provides an essential non-cell-autonomous signal at gastrulation, leading to induction of the floorplate and ventral brain.

Current models for dorsoventral patterning of the vertebrate neural tube involve a signalling cascade that is initiated by notochord-derived Sonic hedgehog (Shh), leading to the sequential induction of the ventral midline floorplate and adjacent motoneurons⁸. Accordingly, the floorplate is absent in *shh* mutant mouse embryos⁹. Nevertheless, under certain conditions, the floorplate can differentiate in the absence of an *shh*-expressing notochord^{10–12}. Conversely, some zebrafish mutants, such as *cyclops* (*cyc*), lack a floorplate despite the presence of an *shh*-expressing notochord¹³.

Studies in several vertebrates have implicated nodal/nodal-related factors in mesoderm induction and patterning^{7,14}, development of the rostral central nervous system⁷ and determination of visceral left–right (L–R) asymmetry¹⁵. We have found that a zebrafish *nodal*-related gene, *znr-1*, corresponds to *cyc* (Fig. 1).

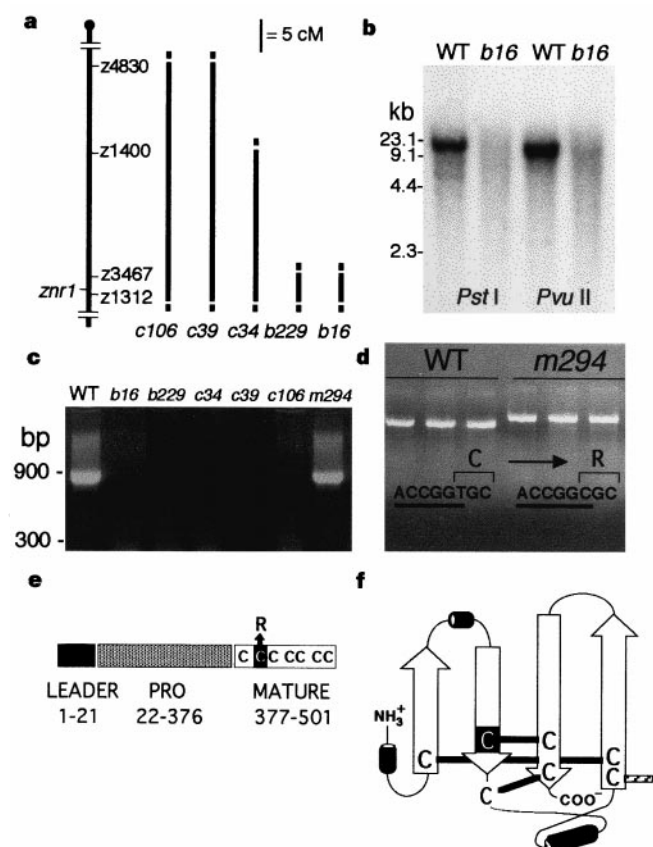


Figure 1 *cyclops* is a *nodal*-related gene. **a**, *znr-1* maps to distal LG 12. Centromere location is approximated and *cyc* deficiencies are indicated (bars). cM, centiMorgans. **b**, Southern blot analysis detects *znr-1* in wild-type (WT) but not *cyc*^{b16} DNA. **c**, *znr-1* primers amplify from WT and *cyc*^{m294} DNA (865 bp) but not from deficiencies. **d**, *Age*I digest of amplified genomic DNA encoding part of the *Znr-1* ligand. The Cys to Arg mutation (bracketed) in *cyc*^{m294} disrupts an *Age* I site (underlined). **e**, Signal, pro and mature ligand domains of Cyc. The mutated Cys is indicated. **f**, Predicted secondary structure; arrows represent β -strands. Conserved cysteines are involved in intramolecular disulphide bonds or ligand dimerization (hatched bar)^{19,20}.

First, typing of a wild-type DNA mapping panel, using the polymerase chain reaction (PCR) and polymorphic markers, localized *znr-1* to linkage group (LG) 12 (Fig. 1a), within the interval where *cyc* was previously mapped^{16,17}. Second, *znr-1* sequences were absent from *cyc* deficiencies (Fig. 1a–c). Third, a *znr-1* polymorphism co-segregated with *cyc* in a haploid DNA mapping panel produced from ethyl nitrosourea (ENU)-induced *cyc*^{m294} (ref. 18) and wild-type sibling embryos. Finally, sequencing of *cyc*^{m294} complementary DNA (confirmed in genomic DNA) revealed a mutation in the second cysteine of the mature ligand region (Fig. 1d–f). Based upon studies of other TGF- β -related ligands^{19,20}, this cysteine is essential for secretion and biological function and this fact, together with genetic data, suggests that *cyc*^{m294} is a null allele. Hereafter, we refer to *znr-1* as *cyclops*.

Expression of *cyc* begins in early gastrulae (30% epiboly; not shown) in cells encircling the blastoderm margin, and is progressively restricted to the dorsal organizer, the embryonic shield (Fig. 2a, b). During mid-gastrulation, expression extends along the midline mesendoderm, including the presumptive prechordal and chorda mesoderm (Fig. 2c). Co-localization of *cyc*, *shh* and *no tail* (*ntl*) expression at this stage (not shown) indicates that *cyc* transcripts are restricted to the mesendodermal layer. However, starting at 90% epiboly, *cyc* transcripts are weakly detected in the ventral neuroectoderm anterior to *pax2* (refs 6, 21) expression at the midbrain–hindbrain boundary (Fig. 2d–f). By the end of gastrulation, *cyc* is expressed in the prechordal mesendoderm and posterior tailbud mesoderm (Fig. 2d–h), but disappears from both regions by the 2–3 somite stage.

Several zebrafish mutations that affect development of axial mesoderm alter *cyc* expression, thereby revealing regulatory interactions in the early embryo. In *cyc*^{m294} (Fig. 2i, j) and *cyc*^{tf291} (not shown) mutants^{18,22} themselves, expression is initially indistinguishable from the wild type, but declines thereafter (Fig. 2i, j), and is absent from the prechordal plate (PCP), a region partly dependent upon *cyc* function^{1,2}. Similarly, *cyc* expression is initially normal in *one-eyed pinhead* (*oep*) embryos, but decreases during gastrulation (Fig. 2k, l), probably contributing to the cyclopia and floorplate defects of *oep* mutants^{18,23}. Thus, function of both *cyc*⁺ and *oep*⁺, which encodes an epidermal growth factor-related protein²⁴, is necessary for maintenance rather than activation of *cyc* expression. However, in severely affected *bozozok*^{m168} (*boz*) mutants¹⁸, dorsal *cyc* expression is greatly reduced or absent throughout gastrulation (Fig. 2m), identifying *boz*⁺ as an early acting, upstream regulator of *cyc*. In contrast to mutations that perturb PCP development, *cyc* expression during gastrulation was not significantly affected by mutations that disrupt notochord development¹² (for example, *floating head* (*flh*) and *ntl*; not shown). An exception was the expanded domain in the *ntl* mutant tail bud (Fig. 2n, o), identifying *Ntl* as a negative regulator of *cyc* expression posteriorly.

A conserved feature of *nodal* genes in mouse, chick and frog is their asymmetric expression in lateral-plate mesoderm (LPM), which is associated with differential left–right morphogenesis (situs) of the embryonic heart and viscera¹⁵. Supporting a similar function during fish embryogenesis, *cyc* is expressed transiently in the left LPM during mid-somitogenesis (Fig. 2p). Moreover, *cyc* is expressed in the left dorsal forebrain (Fig. 2p) in the vicinity of the habenula, a diencephalic region with marked left–right size asymmetry in many vertebrates²⁵. Although mouse *lefty* displays asymmetric expression in the floorplate¹⁵, *cyc* provides the first example of unilateral expression of a TGF- β family member in a discrete domain of the vertebrate brain. Asymmetric *cyc* expression in both the LPM and forebrain becomes bilateral in *flh* and *ntl* mutants (Fig. 2q, r), which is consistent with a role for the notochord in maintaining laterality and with the situs defects reported for these mutants¹⁵. In contrast, LPM expression usually remains left-sided in homozygous *cyc*^{m294} and *cyc*^{tf219} mutants ($n = 20$), which also display heart-situs defects²⁶. Intriguingly, bilateral rather than left-

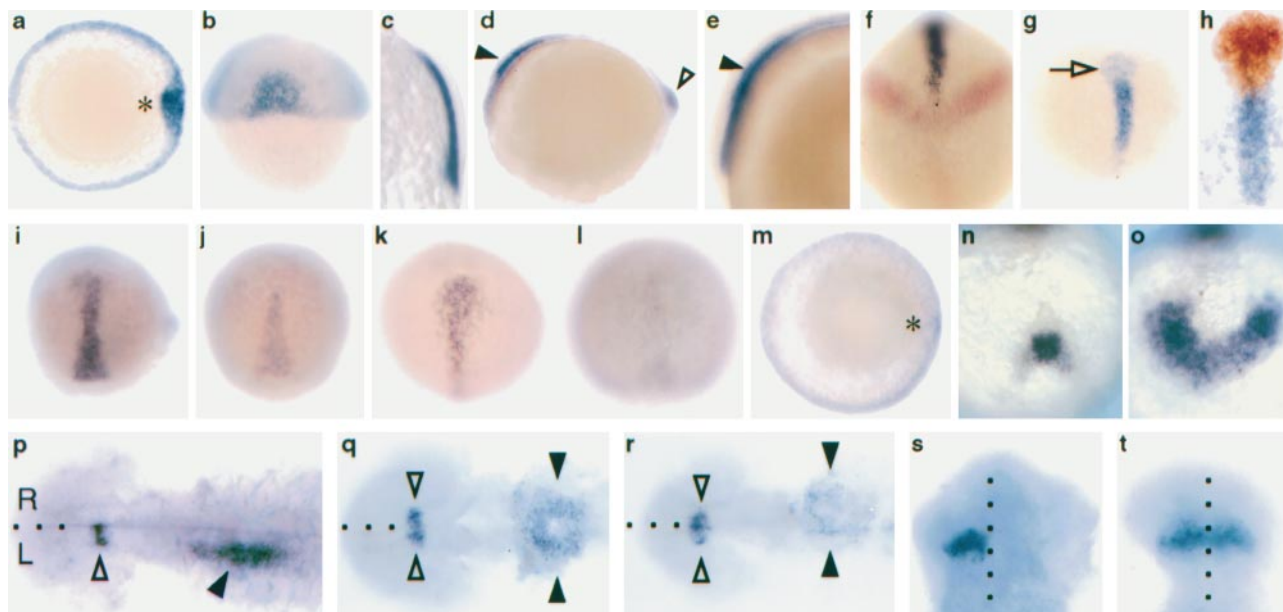


Figure 2 *cyc* expression in wild-type and mutant zebrafish. **a**, 50% epiboly, expression encircling the blastoderm margin and expanded dorsally. **b**, 60% epiboly, expression restricted to dorsal shield. **c**, 50% epiboly, expression in midline mesendoderm. **d-h**, Late gastrula tailbud. *cyc* expression in tailbud mesoderm (**d**, open arrowhead), anteriorly in PCP (**d**, arrowhead), and weakly in ventral brain (**e**). PCP/brain expression is rostral to *pax2* (**f**, pink) at the midbrain-hindbrain boundary^{8,21} and includes the polster (**g**, arrow), which **h**, coexpresses *hgg1* (orange; ref. 1). **i**, **j**, WT (**i**) and *cyc*^{m294} (**j**) at 75% epiboly. **k**, **l**, WT (**k**) and

oep^{m134} (**l**) at 90% epiboly. **m**, 50% epiboly, reduced expression in *boz*^{m168} (compare with **a**). **n**, **o**, WT tailbud expression (**n**) is expanded in *ntl*^{b160} (**o**). **p-t**, 20–22 somite stage. **p**, Asymmetric expression in left (L) LPM (**p**, arrowhead) and forebrain (**p**, open arrowhead), becomes bilateral in *flh*ⁿ¹ (**q**) and *ntl*^{b160} (**r**); midline is indicated. **s**, **t**, Only forebrain expression (**s**) (WT) is bilateral in *cyc*^{m294} (**t**). Views: **a**, **m**, animal pole (* = dorsal); **c**, **d**, **e**, lateral; **n**, **o**, posterior; others, dorsal. Anterior left in **d**, **e**, **p-r**; up in **f**, **g**, **h**, **i-l**, **s**, **t**.

sided expression in the *cyc* mutant forebrain (Fig. 2s, t) reveals that *cyc* regulates its own asymmetric expression within this domain and that visceral and neural left–right specification can be uncoupled.

Mouse and *Xenopus* nodal factors are potent mesoderm inducers^{7,14}. To characterize the inducing activities of *cyc*, we first used the frog animal-cap assay (Fig. 3a, b). *Cyc* consistently induced

mesodermal (*Xbra*, *Xwnt8*, *actin*), organizer (*chordin*) and neural (NCAM) markers. However, unlike the strong mesoderm inducer *Xenopus* nodal-related-2 (*Xnr-2*), *goosecoid* (*gsc*) and *noggin* (not shown), expression was not usually induced by *Cyc*, indicating that the activities of different nodal-related factors are distinguished by this assay. *Cyc*^{m294} RNA induced none of the markers (Fig. 3a), and

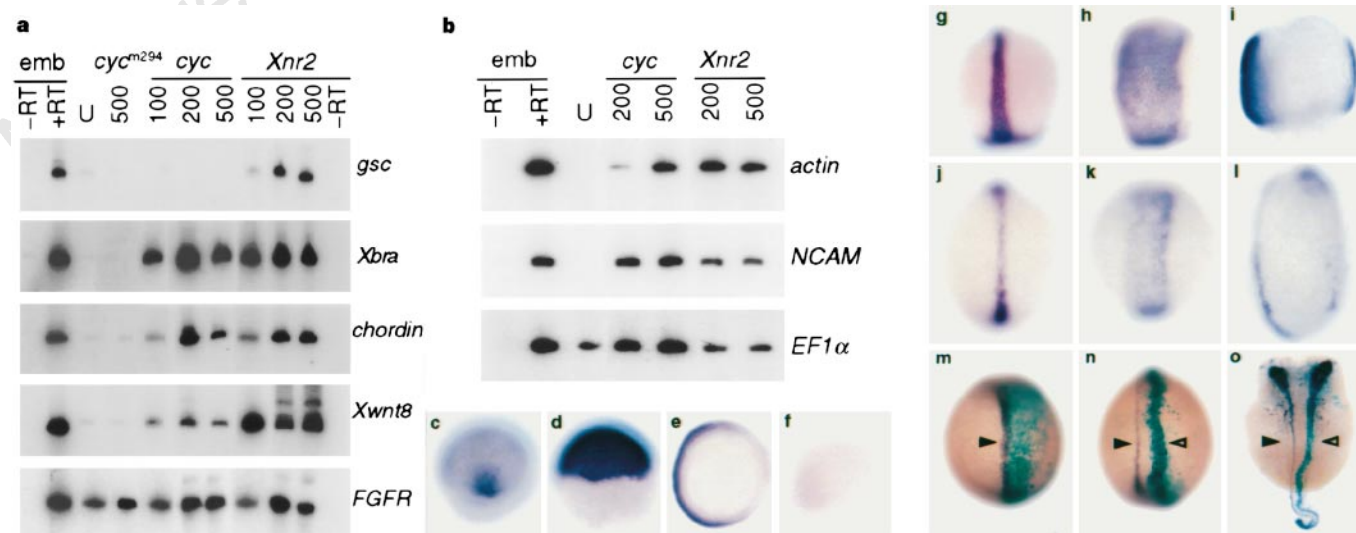


Figure 3 *Cyc* inducing activities. **a**, **b**, Animal caps from frog embryos injected with *cyc*, *Xnr-2* or *cyc*^{m294} RNA (pg per embryo indicated), assayed at early gastrula (**a**) or tailbud (**b**) stages, compared with whole-embryo RNA (emb), uninjected embryo explants (U), or no reverse transcriptase controls (–RT). *FGFR* and *EF1α*: RNA loading controls. **c–f**, Uninjected (**c**, **e**) and *cyc* RNA-injected (**d**, **f**) zebrafish embryos (shield stage), analysed for dorsal marker *goosecoid* (viewed dorsally) (**c**, **d**) and ventrolateral marker *eve1* (viewed anteriorly) (**e**, **f**). **g–n**, Tailbud-

1 somite stage. *ntl* expression in presumptive notochord (**g**) was expanded (**h**) or duplicated (**i**) in *cyc*-injected embryos, as was (**j–l**) *tiggy-winkle hedgehog*, a floor-plate marker²¹. **m**, Lineage-traced cells (β-gal; turquoise) and midline *shh* expression (purple, arrowhead). **n**, Co-injection of β-gal and *cyc* RNAs induced secondary axes. **o**, Such induced axes often had complete *shh* expression and anterior development (26 h).

co-injecting *cyc*⁺:*cyc*^{m294} RNAs (1:1 or 1:5 ratios) did not reduce activity (not shown), confirming the recessive nature of *cyc*^{m294}. Consistent with the effects of mouse *nodal* RNA on zebrafish gastrulation¹⁴, *cyc* misexpression expanded dorsal markers (for example, *gsc*; Fig. 3c, d) at the expense of ventral ones (for example, *eve1*; Fig. 3e, f), and broadened or duplicated the embryonic axis (Fig. 3g–n). The *Cyc*-induced axes (revealed by β -galactosidase lineage tracing) often exhibited complete anterior morphologies, including eyes and ventral forebrain (Fig. 3o). We conclude that *cyc*⁺ encodes a dorsal mesoderm- and neural-inducing factor that is capable of altering cell fate in embryonic tissues.

The inducing activity of *cyc* RNA could also rescue *shh*-expressing cells of the ventral brain and floorplate that are lacking in one-day-old *cyc* mutant embryos¹³ (Fig. 4). Partly rescued *cyc*^{b16} ($n = 15/68$; 22%) or *cyc*^{m294} ($n = 16/42$; 38%) embryos (distinguished by abnormal anterior morphology, and confirmed by PCR genotyping ($n = 16$)), had substantial numbers of *shh*-expressing cells in the ventral brain and neural tube ventral midline (Fig. 4c). Cyclopia was suppressed in two cases (data not shown) in which the ventral forebrain was restored. Presence of a floorplate was further con-

firmed by expression of *F-spondin*, a marker of differentiating floorplate²¹ ($n = 10$; Fig. 4d). Rescue was not observed in control-injected mutants ($n = 0/27$).

Recovery of the ventral neural tube was strongly associated with the dorsal localization of injected *cyc* RNA ($n = 26/27$; 96%), as determined by the position and tissue contribution of descendent lineage-labelled cells (Fig. 4e–g, turquoise cells). However, ectopic overexpression of *cyc* in cells populating the neural tube ventral midline (Fig. 4e, arrowhead) was insufficient to induce an *shh*-positive floorplate, arguing against a direct action of *cyc*⁺ within the neural tube. Colonization of the notochord by *cyc*-overexpressing cells (Fig. 4f, arrowheads) also failed to restore an *shh*-expressing floorplate. Notably, rescue of *shh*-expressing cells within the ventral brain (Fig. 4g, arrow) and neural tube was strongly correlated with the presence of lineage-labelled derivatives of PCP mesoderm (Fig. 4g, arrowheads) including hatching gland cells, eye muscles and head endoderm². Thus, we propose that *cyc* activity originating in the zebrafish organizer, which gives rise to the PCP, is required for correct patterning of the brain and neural tube.

Initially, the defects in *cyc* mutants were thought to arise from an inability of neuroectoderm to respond to induction from underlying mesoderm³. Subsequently, discovery of the reduced axial mesoderm^{1,2}, and rescue of cyclopia and/or brain defects by transplanted wild-type cells at the shield stage⁴, redirected attention to a primary mesodermal deficiency. Together, the spatiotemporal expression pattern, inducing activities and rescue data, lead to the hypothesis that *cyc*⁺ is required non-autonomously at gastrulation, mediating a signalling pathway for ventral brain and floorplate specification that is triggered surprisingly early in zebrafish embryogenesis. This pathway might include *hedgehog* family members. However, because *hh* genes are still expressed in the *cyc* mutant shield^{13,21}, *Cyc*⁺ is not an upstream regulator of their transcription, and may act in a parallel pathway. Unlike *oep*, zebrafish *boz* appears to regulate early *cyc* expression. Furthermore, recent analysis of another zebrafish *nodal*-related gene (C. Erter and C.V.E.W., unpublished data), and its identification as the *cyc*-interacting locus *squint*³¹, reveals an additional level of complexity in nodal signalling during early embryogenesis.

A related early role for *cyc* is in the maintenance and/or proliferation of mesoderm that produces the PCP, a tissue essential for later patterning of the vertebrate forebrain^{27,28} and for separation of the initial eye anlage into bilateral eye fields²⁹. Although mouse *nodal*^{−/−} embryos arrest at gastrulation, chimaeric and genetic interaction studies that allow development to later embryonic stages show that nodal signalling is critical for normal anterior patterning⁷. As with other *nodal* genes, asymmetric LPM expression indicates an additional conserved function for *cyc* in visceral left–right specification. Finally, *cyc* expression in the zebrafish left forebrain establishes a connection between *nodal*-related gene activity and localized brain asymmetry, with implications for the evolution of brain regionalization in higher vertebrates.

Note added in proof: Dawid and colleagues have independently reached similar conclusions in their characterization of *cyclops* and other *nodal*-related genes^{32,33}.

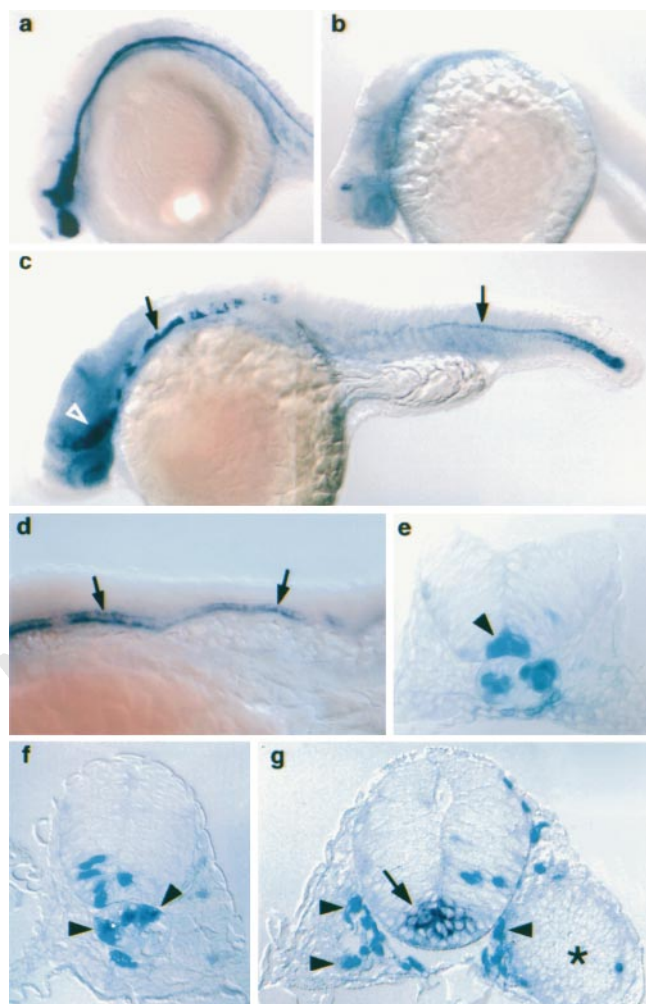


Figure 4 *cyc* RNA rescues ventral neural tube of mutants. **a, b**, Control-injected wild-type (WT) (**a**) with *shh*-expressing ventral brain and floorplate¹³, and *cyc*^{b16} (**b**). At 26 h, *shh* transcripts are sometimes detected in dorsal forebrain, and in posterior-most tail ventral neural tube ($n = 5/19$ *cyc* mutants, 4–10 cells), but not in ventral brain or trunk neural tube. **c, d**, *shh* expression in ventral midbrain (**c**, open arrowhead), hindbrain and trunk floorplate (**c**, arrows) and *F-spondin*-expressing floorplate²¹ (**d**, arrows) in rescued *cyc*^{b16}. **e–g**, Transverse sections (3 μ m) through rescued *cyc*^{m294} (confirmed by genotyping), showing β -gal lineage-labelled (*Cyc*-coexpressing) cells (arrowheads) at the level of hindbrain (**e**), trunk (**f**) and midbrain (**g**), and *shh* expression (arrow).

Methods

Zebrafish stocks. Zebrafish stock maintenance and embryo production were as described previously¹². Gamma ray-induced deletions *cyc*^{b16} and *cyc*^{b229} (ref. 3, 17), *cyc*^{c34}, *cyc*^{c39} and *cyc*^{c106} (our unpublished observations) were maintained in the AB background. ENU-induced alleles *cyc*^{m294} (ref. 18) and *cyc*^{f219} (ref. 22) were maintained in AB and TL wild-type backgrounds, respectively.

znrl1/cyc cDNAs and injection constructs. Partial *znrl1/cyc* cDNAs recovered from a gastrula/neurula zebrafish library by using mouse *nodal* and *Xenopus* *Xnr-1* mature region cDNA probes were used to recover full-length cDNAs from a gastrula library. The *Znr-1* mature ligand is 81%, 78% or 74% similar to mouse *nodal*, *Xenopus* *Xnr-1* or *Xnr-4*, respectively. A full-length *cyc* cDNA was inserted into the RNA synthesis plasmid pCS2⁺ and sense RNA synthesized

from Asp718-linearized plasmid using SP6 RNA polymerase and a commercial kit (Ambion).

Mapping/mutational analyses. *cyc* deficiencies were defined by using Z markers³⁰ and the absence of *znr-1* sequences was confirmed by using intron-flanking primers (5'-GGCCTGCCCGAACCCACTGGGAG-3' and 5'-GGCATCCGCACTCTCCACCTGC-3'). A wild-type DNA map panel of 288 individual haploid progeny of heterozygous AB and DAR¹⁶ wild-type females was assayed by PCR with primers that amplified simple-sequence length polymorphisms³⁰ from LG 12. PCR typing of an independent map panel (96 haploids from female heterozygous for *cyc*^{m294} (AB) and wild-type (DAR) chromosomes) confirmed *cyc/znr-1* cosegregation. The *znr-1* polymorphism was amplified by using primers flanking a (CA)_n repeat in the 3' untranslated region (5'-TGATTCTGAGACACACGAGGAC-3' and 5'-CGTCTGTGAGGG-CATGTGTGTAAG-3'), producing 172 (AB) and ~140 (DAR) base pair (bp) fragments, respectively. Standard Southern blotting analysis used a full-length *znr-1* cDNA probe at high stringency on DNA from pooled 3-day-old deficiency mutant or wild-type siblings. To identify the *cyc*^{m294} lesion, haploids were produced from DAR/AB heterozygous females, and DNA and RNA were extracted from individual gastrulae and used for genotyping as DAR (wild type) or AB (*cyc*), and for reverse transcription of RNA followed by PCR (RT-PCR) with *znr-1* primers. Sequencing of PCR products from two embryos revealed a T-C transition at position 1,291, which was confirmed in *cyc*^{m294} genomic DNA (*n* > 10).

In situ hybridization. Whole-mount *in situ* hybridization was essentially as described previously¹, except that BM purple substrate (Boehringer) was sometimes used. Double-label *in situ* hybridizations were done with digoxigenin- and fluorescein-labelled probes with the appropriate alkaline phosphatase-conjugated antibodies (Boehringer), and visualized as above or with fast red (Boehringer) or 4-iodonitroretazolium violet (Molecular Probes; protocol, C. Sagerstrom and H. Sive, personal communication). Semi-thin sections were prepared as described¹².

Zebrafish embryo injections and lineage tracing. To assess effects in wild-type embryos, *in vitro*-synthesized, capped *cyc* RNA (20 pg) was injected into the yolk of embryos of the 1-4-cell stage. In rescue experiments, *cyc* RNA (5-10 pg) was injected into single blastomeres at the 16-cell stage. Embryos were fixed at 60% epiboly, 10-11 h or 22-24 h post-fertilization for *in situ* hybridization. The localization of *cyc* RNA in injected embryos was monitored at 24 h by epifluorescent detection of green fluorescent protein (GFP; 100 pg) or by beta-galactosidase (β -gal; 50 pg) activity from co-injected RNAs. Control embryos received β -gal and/or GFP RNA at the same doses.

Animal cap assays. *In vitro*-fertilized *Xenopus* embryos at the 1-cell stage were injected with capped *cyc* RNA (wild-type or m294 variant). Animal-cap explants from stage-8 blastulae were cultured under standard conditions until sibling stages 10.5 (gastrula) or 24 (tailbud), and flash-frozen for RNA isolation. RT-PCR for *Xbra*, *gsc*, *Xwnt8*, *chordin*, muscle-specific *actin* (*actin*), *NCAM* and *FGFR* was done by standard methods with primers that have been previously reported (references available upon request).

Received 11 May; accepted 28 July 1998.

1. Thisse, C., Thisse, B., Halpern, M. E. & Postlethwait, J. H. *gooseoid* expression in neuroectoderm and mesoderm is disrupted in zebrafish *cyclops* gastrulas. *Dev. Biol.* **164**, 420-429 (1994).
2. Warga, R. *Origin and Specification of the Endoderm in the Zebrafish*. Thesis, Univ. Tübingen (1996).
3. Hatta, K., Kimmel, C. B., Ho, R. K. & Walker, C. The *cyclops* mutation blocks specification of the floor plate of the zebrafish central nervous system. *Nature* **350**, 339-341 (1991).
4. Hatta, K., Puschel, A. W. & Kimmel, C. B. Midline signaling in the primordium of the zebrafish anterior central nervous system. *Proc. Natl Acad. Sci. USA* **91**, 2061-2065 (1994).
5. Hatta, K. Role of the floor plate in axonal patterning in the zebrafish CNS. *Neuron* **9**, 629-642 (1992).
6. Macdonald, R. et al. Regulatory gene expression boundaries demarcate sites of neuronal differentiation in the embryonic zebrafish forebrain. *Neuron* **13**, 1039-1053 (1994).
7. Varlet, I., Collignon, J., Norris, D. P. & Robertson, E. J. Nodal signaling and axis formation in the mouse. *Cold Spring Har. Symp. Quant. Biol.* **LXII**, 105-113 (1997).
8. Tanabe, Y. & Jessell, T. M. Diversity and pattern in the developing spinal cord. *Science* **274**, 1115-1123 (1996).
9. Chiang, C. et al. Cyclopia and defective axial patterning in mice lacking *Sonic hedgehog* gene function. *Nature* **383**, 407-413 (1996).
10. Pittuello, F., Yamada, G. & Gruss, P. Activin A inhibits Pax-6 expression and perturbs cell differentiation in the developing spinal cord *in vitro*. *Proc. Natl Acad. Sci. USA* **92**, 6952-6956 (1995).
11. Artinger, K. B. & Bronner-Fraser, M. Delayed formation of the floor plate after ablation of the avian notochord. *Neuron* **11**, 1147-1161 (1993).
12. Halpern, M. E. et al. Genetic interactions in zebrafish midline development. *Dev. Biol.* **187**, 154-170 (1997).
13. Krauss, S., Concorde, J. P. & Ingham, P. W. A functionally conserved homolog of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **75**, 1431-1444 (1993).

14. Toyama, R., O'Connell, M. L., Wright, C. V. E., Kuehn, M. R. & Dawid, I. Nodal induces ectopic *gooseoid* and *lim1* expression and axis duplication in zebrafish. *Development* **121**, 383-391 (1995).
15. Varlet, I. & Robertson, E. J. Left-right asymmetry in vertebrates. *Curr. Opin. Genet. Dev.* **7**, 519-523 (1997).
16. Postlethwait, J. H. et al. A genetic linkage map for the zebrafish. *Science* **264**, 699-703 (1994).
17. Talbot, W. S. et al. Genetic analysis of chromosomal rearrangements in the *cyclops* region of the zebrafish genome. *Genetics* **148**, 373-380 (1998).
18. Schier, A. F. et al. Mutations affecting the development of the embryonic zebrafish brain. *Development* **123**, 165-178 (1996).
19. Kingsley, D. M. The TGF- β superfamily: new members, new receptors, and new genetic tests of function in different organisms. *Genes Dev.* **8**, 133-146 (1994).
20. Mason, A. J. Functional analysis of cysteine residues of activin A. *Mol. Endocrinol.* **8**, 325-332 (1994).
21. Ekker, S. C. et al. Patterning activities of vertebrate hedgehog proteins in the developing eye and brain. *Curr. Biol.* **5**, 944-955 (1995).
22. Brand, M. et al. Mutations affecting development of the midline and general body shape during zebrafish embryogenesis. *Development* **123**, 129-142 (1996).
23. Strahle, U. et al. *one-eyed pinhead* is required for development of the ventral midline of the zebrafish (*Danio rerio*) neural tube. *Genes Funct.* **1**, 131-148 (1997).
24. Zhang, J., Talbot, W. S. & Schier, A. F. Positional cloning identifies zebrafish one-eyed pinhead as a permissive EGF-related ligand required during gastrulation. *Cell* **92**, 241-251 (1998).
25. Butler, A. B. & Hodson, W. (eds) *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation* (J. Wiley, New York, 1996).
26. Chen, J.-N. et al. Left-right pattern of cardiac *BMP4* may drive asymmetry of the heart in zebrafish. *Development* **124**, 4373-4382 (1997).
27. Pera, E. M. & Kessel, M. Patterning of the chick forebrain anlage by the prechordal plate. *Development* **124**, 4153-4162 (1997).
28. Dale, J. K. et al. Cooperation of BMP7 and SHH in the induction of forebrain ventral midline cells by prechordal mesoderm. *Cell* **90**, 257-269 (1997).
29. Heisenberg, C. P. & Nusslein-Volhard, C. The function of *silberblick* in the positioning of the eye anlage in the zebrafish embryo. *Dev. Biol.* **184**, 85-94 (1997).
30. Knapik, E. W. et al. A microsatellite genetic linkage map for zebrafish. *Nature Genet.* **18**, 338-343 (1998).
31. Feldman, B. et al. Zebrafish organizer development and germ-layer formation require nodal-related signals. *Nature* (in the press).
32. Rebagliati, M. R., Toyama, R., Fricke, C., Haffter, P. & Dawid, I. B. Zebrafish nodal-related genes are implicated in axial patterning and establishing left-right asymmetry. *Dev. Biol.* **199**, 261-272 (1998).
33. Rebagliati, M. R., Toyama, R., Haffter, P. & Dawid, I. B. *cyclops* encodes a nodal-related factor involved in midline signaling. *Proc. Natl. Acad. Sci. USA* **95**, 9932-9937 (1998).

Acknowledgements. We thank C.-M. Fan, B. Hogan, R. Warga and I. Dawid for discussion; W. Driever, E. Knapik, M. Fishman, C. B. Kimmel, S. Ekker, J. Campos-Ortega, D. Kimelman, D. Wilkinson, D. Turner and C. Nüsslein-Volhard for reagents or zebrafish; S. Fisher, A. Switalski, A. Chilcoat and L. Friedman for map panel DNA; and M. Macurak, A. Pinder, M. Sepanski, D. Lee, B. Cortez, A. Hennessey, M. Ray and K. L. Poon for technical assistance. K.S. and V.K. are supported by NSTB, Singapore; A.L.R., A.M.S.C. and J.O.L. by NIH Training Awards; L.S.K. by the NIH and March of Dimes Birth Defects Foundation; M.E.H. by NSF and the Pew Scholars Program; and C.V.E.W. by the NIH.

Correspondence and requests for materials should be addressed to C.V.E.W. The sequence for *znr1/cyclops* has been deposited in Genbank under accession number U87758.

Apoptosis of CD8⁺ T cells is mediated by macrophages through interaction of HIV gp120 with chemokine receptor CXCR4

Georges Herbein^{*†}, Ulrich Mählke^{*†}, Franak Batliwalla[‡], Peter Gregersen[‡], Todd Pappas[†], John Butler[†], William A. O'Brien[†] & Eric Verdin^{*§}

^{*} The Picower Institute for Medical Research, Manhasset, New York 11010, USA

[†] Department of Medicine, Division of Infectious Diseases, University of Texas Medical Branch, Galveston, Texas 77555, USA

[‡] Division of Biology and Human Genetics, Department of Medicine, North Shore University Hospital/Cornell University Medical College, Manhasset, New York 11030, USA

[§] The Gladstone Institute of Virology and Immunology, PO Box 419100, University of California at San Francisco, California 94141-9100, USA

CD8-positive T cells are thought to play an important role in the control of infection by human immunodeficiency virus (HIV) as a result of their cytotoxic activity and by releasing soluble factors^{1,2}. In AIDS patients, the absolute number of CD8⁺ T lymphocytes is decreased in peripheral blood^{3,4} and their turnover rate is increased, suggesting that there is more cell renewal and cell death occurring⁵. Anti-retroviral therapy raises CD8⁺ T-cell

counts in HIV-infected patients⁶⁻⁸. Here we report that the death rate of CD8⁺ T cells by apoptosis increased markedly during HIV infection of peripheral blood mononuclear cells *in vitro*. Apoptosis is induced in a dose-dependent manner by recombinant envelope glycoprotein gp120 from HIV strain X4, or by stromal-derived factor-1 (SDF-1), the physiological ligand of the chemokine receptor CXCR4. Apoptosis is mediated by the interaction between tumour-necrosis factor- α bound to the membrane of macrophages (mbTNF) and a receptor on CD8⁺ T cells (TNF-receptor II, or TNFRII). The expression of both of these cell-surface proteins is upregulated by HIV infection or by treatment with recombinant gp120 or SDF-1. Apoptosis of CD8⁺ T cells isolated from HIV-infected patients is also mediated by macrophages through the interaction between mbTNF and TNFRII. These results indicate that the increased turnover of CD8⁺ T cells in HIV-infected subjects is mediated by the HIV envelope protein through the CXCR4 chemokine receptor.

We found earlier that HIV infection of peripheral blood mononuclear cells (PBMCs) *in vitro* is associated with a marked increase in apoptosis in both infected and uninfected T lymphocytes, including both CD4⁺ and CD8⁺ populations⁹.

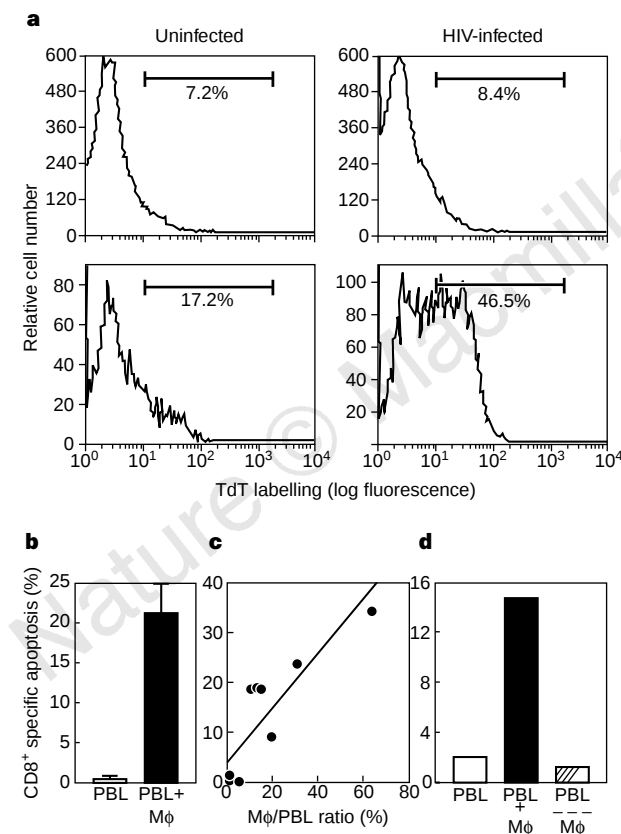


Figure 1 HIV-1 infection enhances macrophage-mediated apoptosis of CD8⁺ T lymphocytes. **a**, Hyperinduction of CD8⁺-T-cell apoptosis after HIV-1 infection of PBL in the presence of macrophages (M ϕ). Apoptosis of CD8⁺ T cells, which were either uninfected or infected with HIV-1_{IIIIB}, at day 9 postinfection was determined by TUNEL assay. One representative experiment of five is shown. **b**, HIV-1_{IIIIB} triggers apoptosis in CD8⁺ T cells in the presence of macrophages. The CD8⁺-T-cell apoptosis is calculated by subtracting the frequency of apoptosis in uninfected cultures from the frequency of apoptosis in infected cultures. Results represent mean values ($\pm$ s.d.) of three independent experiments. **c**, HIV-induced CD8⁺-T-cell apoptosis correlates linearly with the macrophage/PBL ratio. Results represent combined data from five independent experiments. **d**, Measurement of HIV_{89.6}-induced CD8⁺-T-cell apoptosis in purified PBLs (white), PBLs + macrophages (black), or PBLs separated from macrophages by a semipermeable membrane (shaded).

To investigate this effect of HIV infection on CD8⁺-T-cell apoptosis, we infected peripheral blood lymphocytes (PBLs) in the presence or absence of macrophages isolated from the same donor. Addition of macrophages increased apoptosis of CD8⁺ T lymphocytes even in the absence of HIV infection, as previously reported¹⁰ (Fig. 1a)—an effect that was markedly enhanced during HIV infection (Fig. 1a, b). This CD8⁺-T-cell apoptosis increased with the proportion of macrophages present in the co-culture (Fig. 1c), and CD8⁺-T-cell apoptosis during HIV-1 infection occurred in the presence of either autologous or allogenic macrophages, suggesting a non-MHC-restricted mechanism (data not shown). Separation of macrophages from PBLs by a semipermeable membrane blocked HIV-induced apoptosis, indicating that intercellular contact between macrophages and CD8⁺ T cells is necessary for apoptosis induction (Fig. 1d). CD8⁺-T-cell apoptosis was not increased simply as a result of nonspecific macrophage activation, because bacterial lipopolysaccharide (LPS) treatment of the same primary culture system (PBMCs) had no effect on CD8⁺-T-cell apoptosis (data not shown). The functional significance of this increase in apoptosis was illustrated by a 50% reduction in viable CD8⁺ T cells on day 8 after HIV infection (data not shown).

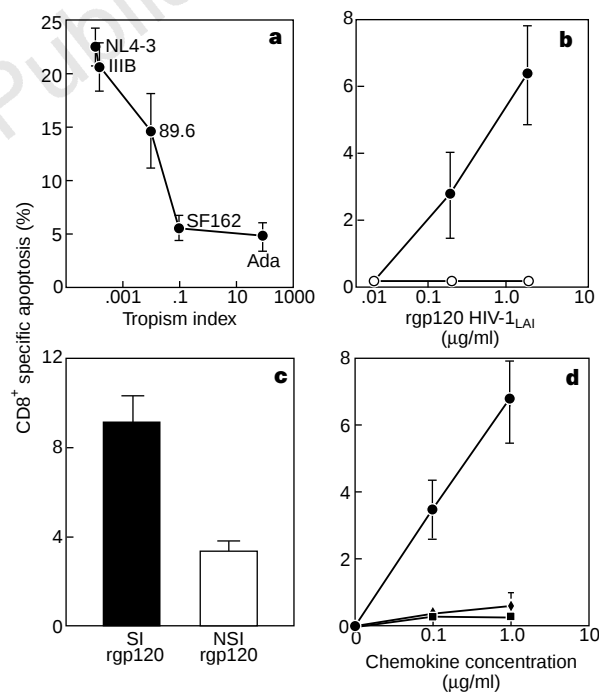


Figure 2 Recombinant gp120 activates macrophages via CXCR4 to kill CD8⁺ T lymphocytes. **a**, Rate of HIV-induced CD8⁺-T-cell apoptosis as a function of viral tropism. CD8⁺-T-cell apoptosis was measured by TUNEL assay after infection with X4 (NL4-3, IIIB), X4R5 (89.6), and R5 (SF162, Ada) HIV-1 strains and is plotted against the tropism index¹¹. **b**, CD8⁺-T-cell apoptosis in response to increasing concentrations of recombinant envelope glycoprotein gp120 (rgp120_{LAI}) measured 18 h after treatment in the presence (filled circles) or absence (circles) of macrophages. **c**, Induction of CD8⁺-T-cell apoptosis by rgp120 (1 μ g ml⁻¹) derived from a syncytium-inducing HIV-1_{CM235} (SI rgp120) and from a non-syncytium-inducing HIV-1_{93TH975} (NSI rgp120) strain, measured by TUNEL assay 18 h after treatment. **d**, CD8⁺-T-cell apoptosis induced in response to rhuSDF-1 (circles), rhuRANTES (diamonds), and rhuMIP-1 β 1 (squares) 18 h after treatment. Results are means ($\pm$ s.d.) of three independent experiments. CD8⁺-T-cell apoptosis did not increase in the absence of macrophages (data not shown). Each data point is the mean ($\pm$ s.d.) of three to five independent experiments.

To determine whether productive macrophage infection is necessary for HIV-induced apoptosis of CD8⁺ T cells, cultures of purified PBLs with or without macrophages were infected with HIV-1 strains with different cellular tropisms. All HIV-1 strains tested induced significant CD8⁺-T-cell apoptosis only in the presence of macrophages (not shown). Specific CD8⁺-T-cell apoptosis followed the peak of viral replication by 48–72 hours. Unexpectedly, the apoptotic response was greater after infection with X4 (NL4-3, HTLV-IIIB) and dual-tropic X4R5 viral strains (89.6) than after infection with R5 strains (SF162, Ada) (Fig. 2a); the macrophage-tropism of a viral strain¹¹ correlated inversely ($r^2 = 0.84$) with its ability to induce CD8⁺-T-cell apoptosis (Fig. 2a).

As HIV cellular tropism is primarily determined by the envelope protein, our results indicated that the HIV envelope glycoprotein gp120 could be a key determinant of CD8⁺-T-cell apoptosis. We therefore tested the ability of purified recombinant envelope protein of HIV-1_{LAI}, an X4 virus, to induce apoptosis and found that there was a macrophage-dependent increase in CD8⁺-T-cell apoptosis in response to increasing amounts of recombinant (r) gp120_{LAI} (Fig. 2b). To confirm that the envelope protein confers the ability to induce CD8⁺-T-cell apoptosis on a viral strain, we tested two

other rgp120 envelope glycoproteins derived from primary HIV-1 isolates, syncytium-inducing (SI) HIV-1_{CM235} or non-syncytium-inducing (NSI) HIV-1_{93TH975}. SI rgp120 induced CD8⁺-T-cell apoptosis three times more effectively than did NSI-rgp120 (Fig. 2c).

A major determinant of HIV cellular tropism is the co-receptor used by the virus to gain entry into its target cells. The chemokine receptors CCR5 and CXCR4 are co-receptors that act with CD4 to permit viral entry and fusion^{12–16}. We therefore tested the role of CXCR4 and CCR5 in CD8⁺-T-cell apoptosis. Stromal-cell derived

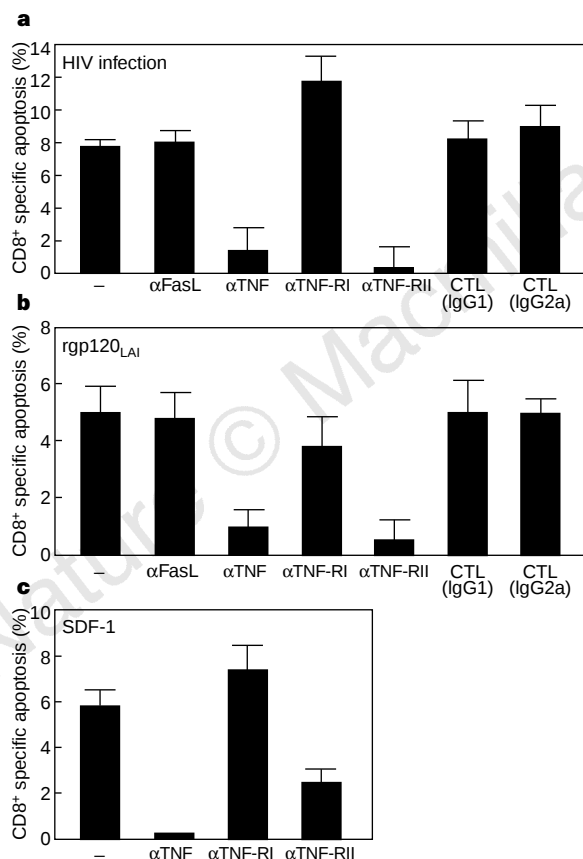


Figure 3 Induction of CD8⁺-T-cell apoptosis during HIV infection or in response to rgp120 and SDF-1 is dependent on interaction between TNF and TNF receptor II. **a**, HIV_{89.6}-infected PBMCs were incubated with neutralizing antibodies against FasL, TNF, TNFRI, TNFRII, or their isotype controls (CTL: IgG₁, IgG_{2a}), and apoptosis of CD8⁺ T cells was measured at day 11 after infection, by TUNEL assay. CD8⁺-specific apoptosis represents the increase in apoptosis measured after HIV infection compared to uninfected controls. **b**, **c**, PBMCs were treated with rgp120_{LAI} (1 μg ml⁻¹) or SDF-1 (1 μg ml⁻¹) for 18 h in the presence of neutralizing antibodies against FasL, TNF, TNFRI, TNFRII, or their isotype controls (IgG₁, IgG_{2a}), and apoptosis of CD8⁺ T cells was measured by TUNEL assay. CD8⁺-specific apoptosis represents the increase in apoptosis measured after treatment with rgp120_{LAI} or SDF-1 compared to untreated controls. Basal CD8⁺ apoptosis varied from 1% to 3% in untreated controls.

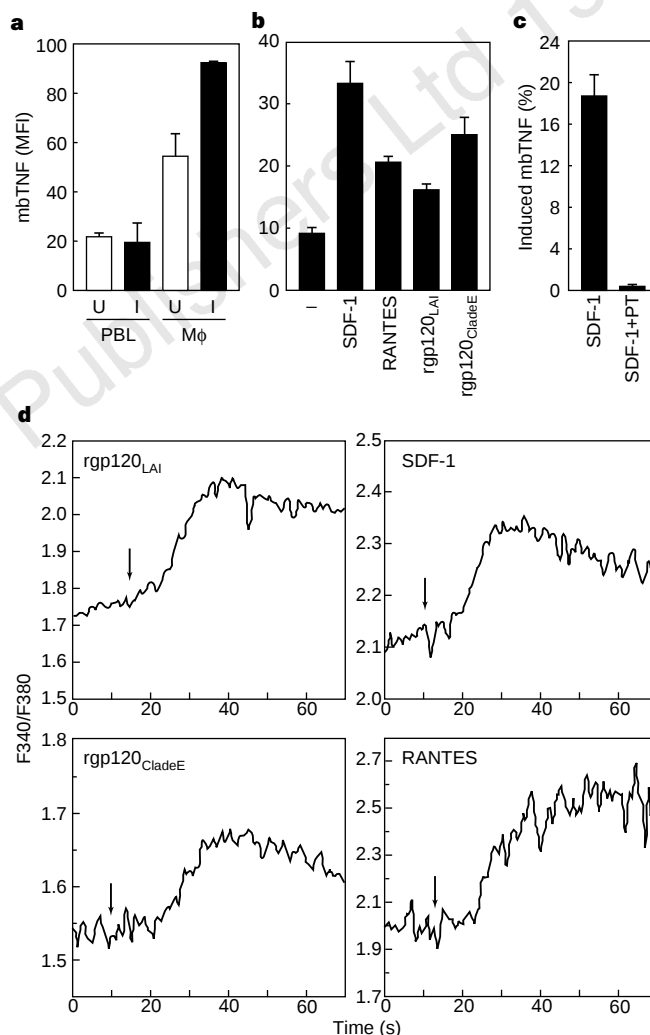


Figure 4 Induction of mbTNF expression in macrophages by HIV infection or envelope is mediated by CXCR4 or CCR5. **a**, Measurement of mbTNF on macrophages and PBLs at day 11 after infection with HIV_{89.6}. mbTNF (mean fluorescence intensity) was measured by flow cytometry on CD14⁺ macrophages and CD45⁺ PBLs after infection of co-cultures of PBL and macrophages; U, uninfected; I, infected. **b**, Measurement of mbTNF expression (mean fluorescence intensity) in response to SDF-1, RANTES, rgp120_{LAI}, rgp120_{CladeE} (NSI isolate, HIV_{93TH975}) (all at 1 μg ml⁻¹). **c**, Induction of mbTNF in macrophages by SDF-1 is inhibited by pertussis toxin. PBMCs were treated with SDF-1 (1 μg ml⁻¹) in the presence or absence of pertussis toxin (200 ng ml⁻¹). mbTNF expression was measured in macrophages by flow cytometry and is expressed as the percentage differential increase between SDF-1-treated and untreated cells. Results represent mean values (±s.d.) of three independent experiments (**a**, **b** and **c**). **d**, Induction of Ca²⁺ flux by SDF-1, RANTES, and rgp120 in macrophages. Fura-2-loaded macrophages were treated with SDF-1, RANTES, or rgp120 (LAI or CladeE) (all at 1 μg ml⁻¹) at the times indicated by the arrows. Results are given as the ratio of fluorescence emissions at 340 and 380 nm (F340/F380) over time. Similar results were obtained in five independent experiments.

factor 1 (SDF-1), a CXCR4 ligand, triggered CD8⁺-T-cell apoptosis in a dose-dependent way, whereas the CCR5 ligands RANTES and MIP-1 β had little or no effect on apoptosis (Fig. 2d). Induction of apoptosis by SDF-1 was dependent on macrophages' being present in the culture (data not shown). As CXCR4 is the receptor used by the envelope X4 viruses to infect T cells, both sets of results indicate that CD8⁺-T-cell apoptosis may be triggered through activation of CXCR4 by X4 HIV strains, by SI rgp120 or by SDF-1.

As intercellular contacts between CD8⁺ T lymphocytes and macrophages are critical for CD8⁺-T-cell apoptosis (Fig. 1d), we investigated the role of cell-surface proteins in the induction of apoptosis in T cells¹⁷. These included Fas and its cognate receptor, Fas ligand (FasL), and TNF- α and its cognate receptors, TNF receptor type I (TNFRI) and type II (TNFRII). HIV-induced CD8⁺-T-cell apoptosis was blocked by neutralizing antibodies against TNF and TNFRII, but not by antibodies against TNFRI or FasL (Fig. 3a). Apoptosis after treatment with rgp120_{LAI} or SDF-1 was also blocked by anti-TNF and anti-TNFRII antibodies, but not by anti-TNFRI or anti-FasL antibodies (Fig. 3b, c), indicating that CD8⁺-T-cell apoptosis is triggered by interaction between macrophage-membrane-bound TNF- α (mbTNF) and TNFRII on CD8⁺ T cells; these results are consistent with the pro-apoptotic role of TNFRII (refs 18, 19).

By using flow-cytometric analysis of macrophages and CD8⁺ T cells, we showed that HIV infection significantly increases mbTNF expression in macrophages but not in PBLs (Fig. 4a). Likewise, SDF-

1, RANTES, rgp120_{LAI}, and rgp120_{CladeE} also induce mbTNF at the surface of primary human macrophages (Fig. 4b). Thus, both CXCR4 and CCR5 appear to be functional in macrophages. In support of this idea, we found that induction of mbTNF expression by SDF-1 was blocked by pertussis toxin, an inhibitor of G-protein-coupled receptors (Fig. 4c). Also, both recombinant HIV envelope proteins and the chemokines RANTES, SDF-1 and MIP-1 β (results not shown) induced Ca²⁺ signalling in primary macrophages, as measured by assay with Fura-2 (ref. 20) (Fig. 4d). Together, these results show that the envelopes of X4 and R5 viruses stimulate expression of macrophage mbTNF by virtue of their interaction with CCR5 and CXCR4 but they do not explain how X4 viruses specifically induce CD8⁺-T-cell apoptosis.

The cell-surface expression of TNFRII and, to a lesser degree, of TNFRI was increased on CD8⁺ T cells, but not on CD4⁺ T cells, during HIV infection (Fig. 5a). TNFRII expression was selectively induced in CD8⁺ T cells, but not in CD4⁺ T cells, by SDF-1 and rgp120_{LAI}, which use CXCR4, but not by RANTES and rgp120_{CladeE}, which use CCR5 (Fig. 5b). In agreement with this, flow-cytometric analysis of the cell-surface expression of CXCR4 and CCR5 on CD8⁺ T cells indicated that CXCR4 expression was significant but that of CCR5 was almost undetectable (Fig. 5c).

These results suggest that the HIV envelope protein activates CXCR4 on macrophages and CD8⁺ T cells, which increases the expression of two cell-surface proteins, mbTNF and TNFRII involved in apoptosis induction. We therefore investigated the

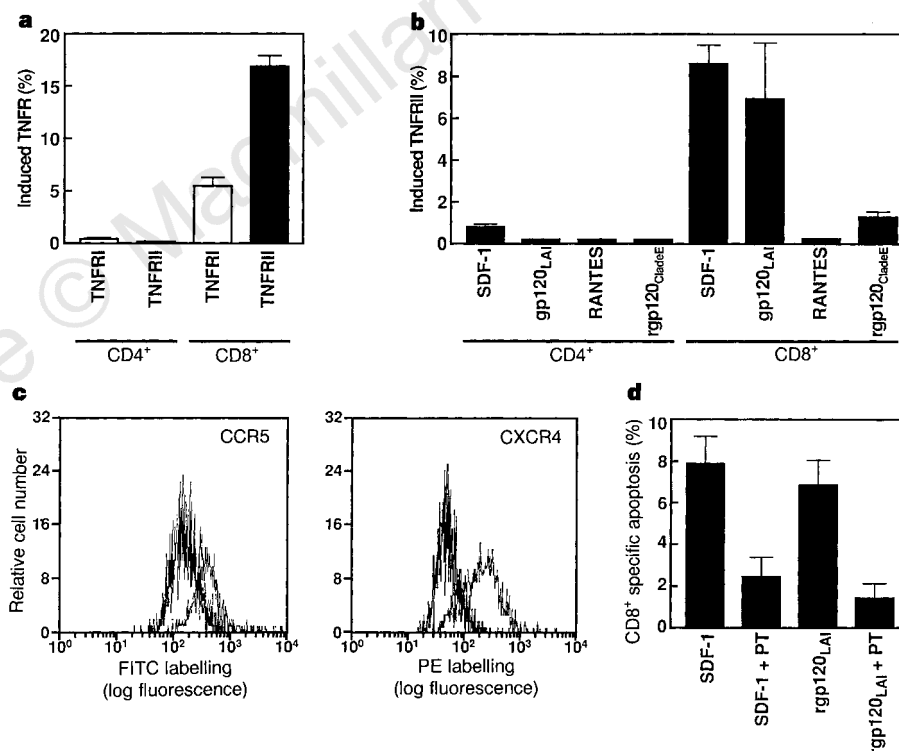


Figure 5 Selective induction of TNF receptor II on CD8⁺ T cells by CXCR4 activation. **a**, Induction of surface TNFRII on CD8⁺ T cells, but not on CD4⁺ T cells, during HIV infection. Surface TNFRI (white bars) and TNFRII (black bars) were measured in CD4⁺ and CD8⁺ populations by flow cytometry at day 11 after infection of co-cultures of PBLs and macrophages with HIV_{99.6}. Results are means ($\pm$ s.d.) of three independent experiments and are expressed as the differential increase in TNFR expression (per cent of positive cells) in HIV-infected cultures in comparison to uninfected cultures. **b**, Induction of cell-surface TNFRII on CD8⁺ T lymphocytes in response to the CXCR4 agonists rgp120_{LAI} and SDF-1. Surface TNFRII was measured in CD4⁺ and CD8⁺ populations by flow cytometric analysis of co-cultures of PBLs and macrophages 18 h after treatment with SDF-1, RANTES, gp120_{LAI} or gp120_{CladeE} (NSI isolate, HIV_{93TH975}). Results are means ($\pm$ s.d.) of three independent experiments and are expressed as the differential increase in TNFR expression (per cent of positive cells) in HIV-infected cultures in comparison to uninfected cultures. **c**, Selective expression of CXCR4 on CD8⁺ T cells. Isolated PBLs were fixed in 3% paraformaldehyde and stained with PerCP-labelled anti-CD8 antibody and either phycoerythrin-labelled anti-CXCR4 or FITC-labelled anti-CCR5. Gated CD8⁺ T cells were analysed for CXCR4 and CCR5 expression by flow cytometry. One representative experiment of two is shown. **d**, Inhibition of SDF-1 and rgp120-induced CD8⁺-T-cell apoptosis by pertussis toxin. PBMCs were treated for 18 h with SDF-1 (1 μ g ml⁻¹) or rgp120_{LAI} (1 μ g ml⁻¹) in the presence or absence of pertussis toxin (PT) (200 ng ml⁻¹), which was added 1 h before SDF-1 and rgp120. Apoptosis of CD8⁺ T cells was measured by TUNEL assay. Results are means ($\pm$ s.d.) from three independent experiments.

s.d.) of three independent experiments and are expressed as the differential increase in TNFR expression (per cent of positive cells) in treated cultures in comparison to untreated cultures. **c**, Selective expression of CXCR4 on CD8⁺ T cells. Isolated PBLs were fixed in 3% paraformaldehyde and stained with PerCP-labelled anti-CD8 antibody and either phycoerythrin-labelled anti-CXCR4 or FITC-labelled anti-CCR5. Gated CD8⁺ T cells were analysed for CXCR4 and CCR5 expression by flow cytometry. One representative experiment of two is shown. **d**, Inhibition of SDF-1 and rgp120-induced CD8⁺-T-cell apoptosis by pertussis toxin. PBMCs were treated for 18 h with SDF-1 (1 μ g ml⁻¹) or rgp120_{LAI} (1 μ g ml⁻¹) in the presence or absence of pertussis toxin (PT) (200 ng ml⁻¹), which was added 1 h before SDF-1 and rgp120. Apoptosis of CD8⁺ T cells was measured by TUNEL assay. Results are means ($\pm$ s.d.) from three independent experiments.

Table 1 Inhibition by antibodies of increased apoptosis of CD8⁺ T cells in patients infected with SI viruses

Patient	CD8 ⁺ T-cell apoptosis (%)							Clinical parameters			
	Untreated	Anti-TNF-RI	Anti-TNF-RII	Anti-TNF	Anti-FasL	Ctl IgG1	Ctl IgG2	CD4	CD8	Clinical status	MT-2
A	1.32	3.05	0.74	0.4	2.92	2.99	1.15	621	1,864	Asym	NSI
B	16.05	19.1	13.91	10.23	16.96	20.72	14.29	394	756	Asym	NSI
C	7.39	9.56	3.53	4.46	6.33	7.9	6.2	510	536	Asym	NSI
D	3.78	3.26	3.2	3.91	4.23	4.1	6.27	106	1,866	AIDS	NSI
E	8.34	5.09	1.17	1.57	5.31	6.39	8.69	8	339	AIDS	NSI
Mean (NSI)	7.4	8.0	4.5	4.1	7.1	8.4	7.3	328	1,072		
F	32.83	17.54	19.14	28.94	27.6	32.5	37.14	48	273	AIDS	SI
G	37.68	28.38	19.75	21.25	33.15	33.78	37.77	21	323	AIDS	SI
H	18.58	17.16	12.23	13.98	15.76	19.18	20.22	138	543	AIDS	SI
I	29.57	33.64	20.37	22.46	27.04	37.83	28.11	118	543	AIDS	SI
Mean (SI)	29.7	24.2	17.9	21.6	25.9	30.8	30.8	81	420		
Mean (NSI + SI)	17.3	15.2	10.4	11.9	15.5	18.4	17.8				

PBMCs were isolated from the blood of HIV-infected subjects and either were left untreated or were treated with 20 $\mu\text{g ml}^{-1}$ anti-TNFR1, 20 $\mu\text{g ml}^{-1}$ anti-TNFR2, 10 $\mu\text{g ml}^{-1}$ anti-TNF, 10 $\mu\text{g ml}^{-1}$ anti-FasL monoclonal antibodies, or isotype controls IgG₁ and IgG_{2a}. The percentage of CD8⁺ T cells undergoing apoptosis 18 h after addition of antibodies was measured by TUNEL assay. CD4 and CD8 levels are represented by counts per mm³ of blood. The predominant viral strain present in each patient was determined by MT-2 assay. Asym, asymptomatic.

effect of pertussis toxin, an inhibitor of signal transduction by chemokine receptors which are coupled to G proteins, and found that, as expected from our results, it inhibited the increase in CD8⁺ T-cell apoptosis induced in response to SDF-1 and gp120_{LAI} (Fig. 5d).

To determine whether this pathogenic mechanism takes place *in vivo*, we used the TUNEL assay to examine PBMCs isolated from HIV-infected subjects for spontaneous CD8⁺ T-cell apoptosis. First, we compared the rate of spontaneous apoptosis in the presence or absence of macrophages. In two patients, the rates of apoptosis were 18.9% and 17.9%, respectively, of the total CD8⁺ T-cell population (normal is 1–4%). Removal of macrophages decreased the rate of CD8⁺ T-cell apoptosis to 10.3% and 8.0%, respectively, demonstrating the importance of macrophages in mediating CD8⁺ T-cell apoptosis in HIV-infected subjects.

Next, we analysed spontaneous CD8⁺ T-cell apoptosis and the effect of blocking antibodies on this process in HIV-infected subjects (Table 1). Patients infected with SI HIV strains, as determined by MT-2 assay for X4 viral strains, had a much higher rate of CD8⁺ T-cell apoptosis than patients infected with NSI viruses (29.7% versus 7.4%). Patients infected by SI isolates also had lower CD8⁺ T-cell counts than patients infected by NSI viruses (420 versus 1,072 cells per mm³). These results agree with our *in vitro* observations that X4 HIV strains induce more CD8⁺ T-cell apoptosis than do R5 strains. However, patients infected with SI HIV strains have a different immunological profile, with marked immune activation, and additional factors may operate to increase CD8⁺ T-cell apoptosis in these patients. In the presence of neutralizing antibodies against TNFR1, TNFR2, TNF- α , FasL or control isotypes for these antibodies, CD8⁺ T-cell apoptosis was inhibited primarily by anti-TNFR2 (10.4% versus 17.8% for the isotype control) and by anti-TNF (11.9% versus 18.4% for the isotype control antibody) (Table 1). As these two antibodies also inhibited HIV-induced CD8⁺ T-cell apoptosis *in vitro* (Fig. 3a), our results suggest that the same mechanism causes CD8⁺ T-cell apoptosis in HIV-infected subjects. As antibody treatment did not completely suppress apoptosis, it may already have been initiated in the infected patients and so was not inhibited *in vitro*.

Our results show that stimulation of macrophages by X4 HIV envelope protein leads to increased apoptosis of CD8⁺ T cells. SI HIV-1 strains can be recovered in almost half of HIV-infected subjects, usually late in the course of infection, and are associated with accelerated clinical progression^{21–24}. Therefore, SI/X4 viruses appear to exert their deleterious effect on the immune system not only by direct cytopathic effects on CD4⁺ T cells, but also by indirect killing of CD8⁺ T cells. Because CD8⁺ T cells are thought to limit the spread of the infection either by their cytotoxic activity or by the

release of soluble factors, the failure of CD8⁺ T-lymphocyte function could contribute to the development of AIDS^{1,2,25,26}. □

Methods

Isolation and culture of peripheral blood lymphocytes and blood monocyte-derived macrophages. Human PBMCs were isolated from healthy donors as described⁹. Adherent tissue culture differentiated macrophages (>94% CD14⁺ by flow cytometric analysis) were cultured in RPMI medium supplemented with 10% (v/v) pooled AB human serum (Sigma). Non-adherent cells (>95% PBLs, as assessed by CD45⁺, Simultest Leucogate, Becton Dickinson) were grown in RPMI with 10% FCS supplemented for the first 48 h with phytohaemagglutinin A (PHA, 5 $\mu\text{g ml}^{-1}$; Sigma) before addition of human recombinant interleukin-2 (hrIL-2, 20 IU ml⁻¹; Life Technologies). For co-culture, PHA/IL-2-activated PBLs were mixed with 10% macrophages and grown in RPMI with 10% (v/v) FCS supplemented with hrIL-2 as described above. To block PBL-macrophage intercellular contacts, the cell populations were separated by a semipermeable membrane in six-well plates (0.4- μm pore size; Transwell, Costar).

Generation of viral stocks. HIV-1 strains NL4-3, HTLV-IIIB, 89.6, SF162 and Ada were provided by the AIDS Research and Reference Reagent Program (NIAID, NIH). Viral stocks were prepared from supernatants after filtration through a 0.45- μm filter, quantified by measurement of reverse transcriptase (RT) activity, and stored at -70 °C.

Infections. Three-day-old PHA/IL-2-activated PBL or macrophage-PBL co-cultures were infected with virus (500,000 c.p.m. RT ml⁻¹). Culture supernatants were collected every 2–3 days and assayed for RT activity. Cell pellets were collected, washed twice with PBS, and fixed with 3% paraformaldehyde in PBS for 30 min. Fixed cells were analysed by flow cytometry.

Recombinant gp120 treatment. Three-day-old PHA/IL-2-activated PBL or macrophage-PBL co-cultures were treated with rgp120 from LAI, CM235 or 93TH975 HIV-1 strains (NIH AIDS Research and Reference Reagent Program). At 18 h after treatment, cells were collected, washed, fixed and analysed by flow cytometry as above.

Chemokine treatment. Three-day-old PHA/IL-2-activated PBL or macrophage-PBL co-cultures were treated with rhSDF-1, rhRANTES, or rhMIP-1 β (R&D Systems). At 18 h after treatment, cells were collected, washed, fixed and analysed as above.

Antibody treatment. Macrophage-PBL co-cultures were treated with 20 $\mu\text{g ml}^{-1}$ anti-huTNFR1 mouse IgG₁ (R&D Systems), 20 $\mu\text{g ml}^{-1}$ anti-huTNFR2 mouse IgG_{2a} (R&D Systems), 10 $\mu\text{g ml}^{-1}$ anti-huFas ligand mouse IgG_{2a} (NOK-2; Pharmingen), 10 $\mu\text{g ml}^{-1}$ anti-hu TNF mouse IgG₁ (clone 1825; R&D Systems), or IgG₁ and IgG_{2a} mouse isotype controls (R&D Systems) for 24 h before and during the experiment. Culture supernatants and cells were collected at various times after infection.

Flow-cytometry analysis. Cells fixed in 3% paraformaldehyde in PBS were labelled for TUNEL assay or for FasL, mbTNF, TNFR1 or TNFR2 detection, and analysed with a FACScan flow cytometer (Becton Dickinson) as reported⁹.

CD8 was detected using PercP-labelled anti-human CD8 mouse IgG₁ monoclonal antibodies (Becton Dickinson). Data from 5×10^4 cells were collected, stored and analysed by using Cellquest software (Becton Dickinson). For detection of mbTNF and expression of surface TNFRI and TNFRII, cells were stained with a phycoerythrin-labelled anti-huTNF mouse IgG₁ monoclonal antibody (Pharmingen), FITC-labelled anti-huTNFRI goat polyclonal IgG, FITC-labelled anti-huTNFRII goat polyclonal IgG antibodies, or their respective isotype controls (all from Santa Cruz Biotechnology). CXCR4 and CCR5 expression was measured on CD8⁺ cells by double-labelling with PercP-labelled anti-human CD8 mouse IgG₁ monoclonal antibody (Becton Dickinson) and either mouse monoclonal anti-huCXCR4 (clone 12G5, IgG_{2a}; R&D Systems) or mouse monoclonal anti-huCCR5 (clone 45531.111, IgG_{2b}; R&D Systems).

Terminal transferase dUTP nick end-labelling assay (TUNEL assay). After fixation with 3% paraformaldehyde in PBS for 30 min, 5×10^6 cells per sample were washed three times with 0.3% (v/v) Triton X-100 in PBS, and a TUNEL assay was performed⁹. Apoptosis was measured in both uninfected and infected cells isolated from the same donor, and specific apoptosis in the CD8⁺ T-cell subpopulation was expressed as a percentage of apoptotic cells in the infected culture minus the percentage of apoptotic cells in the uninfected culture.

Fura assay for Ca²⁺ flux determination. Macrophages were scraped from tissue-culture flasks in the presence of 5 mM EGTA/PBS at 4°C. Cells were washed in PBS and loaded with Fura-2/AM (Calbiochem) in HEPES-buffered saline (HBS; 137 mM NaCl, 3 mM KCl, 1.2 mM MgCl₂, 1.8 mM CaCl₂, 10 mM glucose, 10 mM HEPES, pH 7.4) at room temperature for 30 min. Cells were washed once in HBS to remove excess Fura-2/AM. Fluorescence emission (at 510 nm) from excitation at 340 and 380 nm was measured in 1-ml suspensions of 10^6 cells ml⁻¹ with a model F-2000 fluorescence spectrophotometer (Hitachi Instruments). Using published calibration equations, we found that resting Ca²⁺ was at or below 150 nM (not shown). Chemokines and rgp120 proteins were added as 100× solutions in water.

Patients. PBMCs from HIV-infected subjects were isolated by Ficoll-Hypaque as reported⁸. PBMCs were distributed in six-well plates at 5×10^6 cells per well in RPMI supplemented with 10% FCS. At 24 h after isolation, PBMCs were left untreated or were treated with 20 µg ml⁻¹ anti-huTNFRI mouse IgG₁, 20 µg ml⁻¹ anti-huTNFRII mouse IgG_{2a}, 10 µg ml⁻¹ anti-huTNF mouse IgG₁ (clone 1825), 10 µg ml⁻¹ anti-human FasL mouse IgG_{2a} (NOK-2; Pharmingen), and the respective isotype controls (all from R&D Systems). At 18 h after treatment, cells were fixed in 3% paraformaldehyde and analysed for apoptosis by using the TUNEL assay. The CD8⁺ T-cell subpopulation was characterized by using an anti-huCD8 PerCP-labelled monoclonal antibody.

Received 22 June; accepted 3 August 1998.

- Cocchi, F. *et al.* Identification of RANTES, MIP-1α and MIP-1β as the major HIV-suppressive factors produced by CD8⁺ T cells. *Science* **270**, 1811–1815 (1995).
- Levy, J. A., Mackewicz, C. E. & Barker, E. Controlling HIV pathogenesis: the role of the noncytotoxic anti-HIV response of CD8⁺ T cells. *Immunol. Today* **17**, 217–224 (1996).
- Meyaard, L. *et al.* Programmed death of T cells in HIV-1 infection. *Science* **257**, 217–219 (1992).
- Margolick, J. B. *et al.* Failure of T-cell homeostasis preceding AIDS in HIV-1 infection. The Multicenter AIDS Cohort Study. *Nature Med.* **1**, 674–680 (1995).
- Wolthers, K. C. *et al.* T cell telomere length in HIV-1 infection: no evidence for increased CD4⁺ T cell turnover. *Science* **274**, 1543–1547 (1996).
- Mackewicz, C. E., Landay, A., Hollander, H. & Levy, J. A. Effect of zidovudine therapy on CD8⁺ T cell anti-HIV activity. *Clin. Immunol. Immunopathol.* **73**, 80–87 (1994).
- Sarin, A. *et al.* Inhibition of activation-induced programmed cell death and restoration of defective immune responses of HIV⁺ donors by cysteine protease inhibitors. *J. Immunol.* **153**, 862–872 (1994).
- Pakker, N. G. *et al.* Biphasic kinetics of peripheral blood T cells after triple combination therapy in HIV-1 infection: a composite of redistribution and proliferation. *Nature Med.* **4**, 208–214 (1998).
- Herbein, G., Van Lint, C., Lovett, J. L. & Verdin, E. Distinct mechanisms trigger apoptosis in human immunodeficiency virus type 1-infected and in uninfected bystander T lymphocytes. *J. Virol.* **72**, 660–670 (1998).
- Wu, M. X., Daley, J. F., Rasmussen, R. A. & Schlossman, S. F. Monocytes are required to prime peripheral blood T cells to undergo apoptosis. *Proc. Natl Acad. Sci. USA* **92**, 1525–1529 (1995).
- Collin, M., Illei, P., James, W. & Gordon, S. Definition of the range and distribution of human immunodeficiency virus macrophage tropism using PCR-based infectivity measurements. *J. Gen. Virol.* **75**, 1597–1603 (1994).
- Simmons, G. *et al.* Primary, syncytium-inducing human immunodeficiency virus type 1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *J. Virol.* **70**, 8355–8360 (1996).
- Feng, Y., Broder, C. C., Kennedy, P. E. & Berger, E. A. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **272**, 872–877 (1996).
- Bleul, C. C. *et al.* The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* **382**, 829–833 (1996).
- Choe, H. *et al.* The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85**, 1135–1148 (1996).
- Doranz, B. J. *et al.* A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* **85**, 1149–1158 (1996).

- Effros, R. B. & Pawelec, G. Replicative senescence of T cells: does the Hayflick Limit lead to immune exhaustion? *Immunol. Today* **18**, 450–454 (1997).
- Heller, R. A., Song, K., Fan, N. & Chang, D. J. The p70 tumor necrosis factor receptor mediates cytotoxicity. *Cell* **70**, 47–56 (1992).
- Zheng, L. *et al.* Induction of apoptosis in mature T cells by tumour-necrosis factor. *Nature* **377**, 348–351 (1995).
- Gryniewicz, G., Poenie, M. & Tsien, R. Y. A new generation of Ca²⁺ indicators with greatly improved fluorescence properties. *J. Biol. Chem.* **260**, 3440–3450 (1985).
- Schuitmaker, H. *et al.* Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. *J. Virol.* **66**, 1354–1360 (1992).
- Koot, M. *et al.* Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4⁺ cell depletion and progression to AIDS. *Ann. Intern. Med.* **118**, 681–688 (1993).
- Daar, E. S. *et al.* Sequential determination of viral load and phenotype in human immunodeficiency virus type 1 infection. *AIDS Res. Hum. Retrovir.* **11**, 3–9 (1995).
- Connor, R. I. & Ho, D. D. Human immunodeficiency virus type 1 variants with increased replicative capacity develop during the asymptomatic stage before disease progression. *J. Virol.* **68**, 4400–4408 (1994).
- Phillips, R. E. *et al.* Human immunodeficiency virus genetic variation that can escape cytotoxic T-cell recognition. *Nature* **354**, 453–459 (1991).
- Feinberg, M. B. & McLean, A. R. AIDS: decline and fall of immune surveillance? *Curr. Biol.* **7**, R136–R140 (1997).

Acknowledgements. We thank C. Amella for help with RT assays; D. Schepp for blood from HIV-infected patients; M. Grell, J. Nichols, N. Roberts, H. Schmidmayrova and B. Sherry for discussion; N. Shea for graphics; G. Howard and S. Ordway for editing the manuscript; and the AIDS Research and Reference Reagent Program (NIAID, NIH, PHS) for reagents. This work was supported in part by grants from the NIH of the US Public Health Service (to E.V.), by institutional funds from the Picower Institute for Medical Research, the University of Texas Medical Branch, and the Gladstone Institute of Virology and Immunology.

Correspondence and requests for materials should be addressed to E.V. (e-mail: everdin@gladstone.ucsf.edu).

The p35/Cdk5 kinase is a neuron-specific Rac effector that inhibits Pak1 activity

Margareta Nikolic*, Margaret M. Chou†§, Wange Lu‡, Bruce J. Mayer‡ & Li-Huei Tsai*

* Howard Hughes Medical Institute and Department of Pathology,

† Department of Cell Biology, and ‡ Howard Hughes Medical Institute and Children's Hospital and Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

Cyclin-dependent kinase 5 (Cdk5) and its neuron-specific regulator p35 (refs 1–4) are essential for neuronal migration and for the laminar configuration of the cerebral cortex^{5–7}. In addition, p35/Cdk5 kinase concentrates at the leading edges of axonal growth cones and regulates neurite outgrowth in cortical neurons in culture⁸. The Rho family of small GTPases is implicated in a range of cellular functions, including cell migration and neurite outgrowth^{9–13}. Here we show that the p35/Cdk5 kinase co-localizes with Rac in neuronal growth cones. Furthermore, p35 associates directly with Rac in a GTP-dependent manner. Another Rac effector, Pak1 kinase^{14,15}, is also present in the Rac–p35/Cdk5 complexes and co-localizes with p35/Cdk5 and Rac at neuronal peripheries. The active p35/Cdk5 kinase causes Pak1 hyperphosphorylation in a Rac-dependent manner, which results in down-regulation of Pak1 kinase activity. Because the Rho family of GTPases and the Pak kinases are implicated in actin polymerization^{16–18}, the modification of Pak1, imposed by the p35/Cdk5 kinase, is likely to have an impact on the dynamics of the reorganization of the actin cytoskeleton in neurons, thus promoting neuronal migration and neurite outgrowth.

The p35/Cdk5 kinase and the Rho family of GTPases induce shape changes in fibroblasts (ref. 17 and unpublished data) and have been implicated in neurite outgrowth, axon guidance and cell migration^{9–13}. To explore a possible interaction between the p35/Cdk5 kinase and the Rho family of GTPases, we expressed p35 in

§ Present address: University of Pennsylvania, School of Medicine, Department of Cell and Developmental Biology, 36 Street, Hamilton Walk, Philadelphia, Pennsylvania 19104, USA.

COS-7 cells with either active (V12 or V14, GTP-bound) or inactive (N17, GDP-bound) mutants of Cdc42, Rac or Rho fused to glutathione *S*-transferase (GST). A GTP-dependent association was found between p35 and GST–RacV12 (Fig. 1a) as no detectable p35 bound to GST–RacN17. To determine whether the Rac–p35 complexes also contain Cdk5, a dominant-negative mutant of Cdk5 (Cdk5N144 (ref. 19)) was used, which binds more stably with p35 (ref. 8). In the presence of p35, Cdk5N144 also associated with active RacV12 (Fig. 1b). To verify the binding specificity between p35 and Rac, p35 was coexpressed with myc-tagged RacL61 or with two mutants of the Rac effector domain, RacL61A37 and RacL61K40. RacL61A37 was shown to bind the Pak kinases but had lost its ability to induce lamellipodia, whereas RacL61K40 was shown to induce lamellipodia but no longer associated with the Pak kinases^{20,21}. Although a clear association was seen between p35 and RacL61, both the L61A37 and L61K40 Rac mutants lost their ability to bind to p35 (Fig. 1c).

To determine whether the p35–Rac interactions are direct, bacterially produced RacL61 and RacL61K40 were labelled with [³²Pγ-GTP], incubated with equal amounts of GST–p35, GST–Cdk5 or GST and examined by Cerenkov counting. On average, a 3.5-fold-stronger binding between [³²Pγ-GTP]–RacL61 and GST–p35 than [³²Pγ-GTP]–RacL61 with GST–Cdk5 or GST was seen (Fig. 1d). The RacL61K40 mutant failed to interact with GST–p35. A Pak1 mutant (GST–Pak1R299) was used as a control and bound

RacL61 but not RacL61K40, as expected. Together, these observations indicate a direct and specific physical association between p35 and Rac in a GTP-dependent manner, which indicates that the p35/Cdk5 kinase is a downstream target of Rac.

Active Rac has been shown to localize to plasma membranes by carboxy-terminal geranyl-geranylation²². To determine the subcellular distribution of p35, lysates obtained from cortices at embryonic day 18 (E18) were fractionated into membrane, cytoplasmic and nuclear pools and analysed by western blotting. The bulk of p35 was associated with the membrane (Fig. 2A), whereas most cdk5 was present in the cytoplasm with a smaller proportion at the membranes. Negligible amounts of p35 and Cdk5 were seen in the nucleus. As a control, the same extracts were examined for the distribution of Rac and SV2, a neuronal membrane marker. To examine if neuronal p35 and Rac associate, p35 immunoprecipitations were carried out from cortical lysates of p35^{+/–} and p35^{–/–} mice. A strong interaction was seen between p35 and Rac in lysates obtained from p35^{+/–} cortices, whereas no association was evident in p35^{–/–} cortices (Fig. 2B). In confirmation, p35/Cdk5 association was only seen in p35^{+/–} cortices. To investigate whether p35 and Rac co-localize, we initially examined the subcellular distribution of transfected p35 and Rac in Swiss3T3 fibroblasts by immunocytochemistry. Co-localization of RacV12 and p35 was evident particularly in the most actively ruffling areas and at the peripheries of these cells (see arrows in Fig. 2Ca). In primary neurons, co-localization of p35 and Rac was evident in lamellipodia-rich areas of axonal growth cones (Fig. 2Cc) where Cdk5 was concentrated (Fig. 2Cd). p35 and Rac also co-localized in neuronal soma (Fig. 2Cb). Together, these studies establish that the p35/Cdk5 kinase is likely to be a neuron-specific effector of Rac.

Because Rac is known to have many downstream effectors, we investigated whether the p35/Cdk5 kinase is part of a previously identified pathway. Among the Rac effectors, the Pak kinase family has been extensively characterized^{14,15}, two members of which, Pak1 and Pak3 (Pakα and β in rats), are highly enriched in brain^{23,24}. Upon association with GTP–Rac or Cdc42, the Pak kinases auto-phosphorylate and become active. When p35 was coexpressed with myc-tagged Pak1 and RacL61, association between p35 and Pak1 was detected, whereas no interactions were evident in the absence of active Rac (Fig. 2D). Coexpression of RacL61K40, Pak1 and p35 did not facilitate trimolecular complex formation (data not shown), which further suggests that the p35–Pak1 association is dependent on active Rac. To determine whether p35/Cdk5, Rac and Pak1 form complexes in neurons, Pak1 immunoprecipitations were done using cortical lysates obtained from wild-type, p35^{+/–} or p35^{–/–} mice, revealing co-immunoprecipitated Cdk5 from wild type, less from p35^{+/–} and negligible amounts from p35^{–/–} cortices (Fig. 2E). The levels of Pak1–Rac association were unaltered. In primary neurons, p35 and Cdk5 co-localize in axonal growth cones⁸. To examine the localization of Pak1, primary neurons were co-immunostained for Pak1 and Cdk5 or Pak1 and Rac. Prominent co-localization of Cdk5 and Pak1 was evident in axonal growth cones (Fig. 2Ce) where, as expected, extensive Pak1 and Rac co-localization was also seen (Fig. 2Cf). Together, these experiments demonstrate that Rac can associate simultaneously with both p35/Cdk5 and Pak1 kinases and that this complex is present in the highly dynamic areas of neurons such as the growing tips of axons.

Upon coexpression with p35, Cdk5 and active Rac, slower mobility species of Pak1 were apparent; these became more prominent with increasing amounts of the p35/Cdk5 kinase (Fig. 3a). The slower-migrating forms of Pak1 were abolished in the absence of active Rac or in the presence of RacL61K40, and they were greatly reduced upon coexpression with Cdk5N144. Treatment of Pak1 immunoprecipitates with phosphatase diminished the slower-migrating species, which showed that these were phosphorylated forms of Pak1 (Fig. 3b). These experiments indicate that the active p35/Cdk5 kinase can induce hyperphosphorylation of Pak1. Densi-

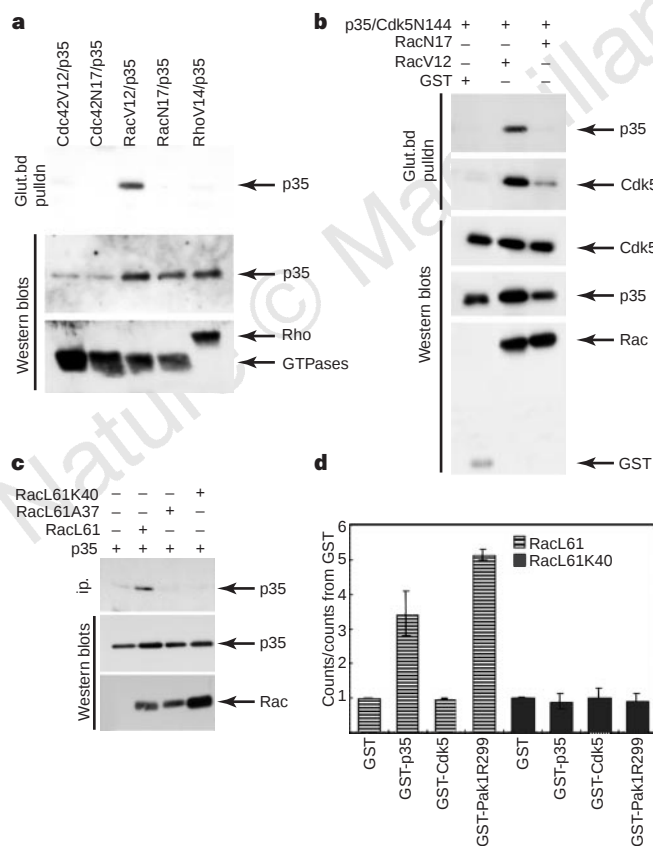


Figure 1 The p35/Cdk5 kinase associates with Rac in a GTP-dependent manner. **a**, p35 western blot of glutathione-bead pulldowns from COS-7 cell lysates expressing p35 and the GTPase mutants (as shown) fused to GST. **b**, As in **a**, including Cdk5N144. **c**, Association between p35, myc–RacL61, myc–RacL61K40 or myc–RacL61A37 was examined in expressing COS-7 cells by immunoprecipitation with antibody 9E10 followed by p35 western blot analysis; lower panels in **a–c** show protein expression levels. **d**, GST–p35, GST–Cdk5, GST–Pak1R299 or GST bound to glutathione beads were incubated with [³²Pγ-GTP]–RacL61 or [³²Pγ-GTP]–RacL61K40 and examined by Cerenkov counting.

toxic analysis showed that high concentrations of the p35/Cdk5 kinase result in phosphorylation of ~75% of total Pak1 expressed, indicating that a large proportion of this protein had associated with the p35/Cdk5 kinase. To examine whether the phosphorylated forms of Pak1 were due to autophosphorylation or phosphorylation by the p35/Cdk5 kinase, we used a catalytically inactive mutant of Pak1 (ATP binding site)^{16,25}. Inactive Pak1 displayed the same mobility shifts as wild-type Pak1 (Fig. 3c), indicating direct phosphorylation by the p35/Cdk5 kinase. To test whether Pak1 is a substrate of the p35/Cdk5 kinase, purified GST–Pak1R299 was used as an *in vitro* substrate for the recombinant GST–p25/GST–Cdk5 kinase. p25 is an amino-terminally truncated form of p35 that is able to interact with and activate Cdk5 (ref. 1). GST–Pak1R299 was phosphorylated by the p25/Cdk5 kinase whereas no phosphorylation of GST was observed (Fig. 3d). Inhibition of p25/Cdk5 with a specific inhibitor, roscovitine²⁶, caused the loss of GST–Pak1R299 phosphorylated species.

To compare GST–Pak1R299 phosphorylated in transfected COS-7 cells in the presence of the p35/Cdk5 kinase and RacL61 with GST–Pak1R299 phosphorylated *in vitro* by the p25/Cdk5 kinase, we performed two-dimensional phospho-tryptic peptide analysis. Six

distinct GST–Pak1R299 peptides were apparent *in vivo*, whereas the *in vitro* phosphorylated GST–Pak1R299 resolved into one major peptide that perfectly comigrated with one of the strongest-labelled peptides *in vivo* (Fig. 3e, peptide X). Upon longer exposure, another labelled peptide was apparent that also had a labelled counterpart *in vivo* (Fig. 3e, peptide Y). Therefore, it is likely that Pak1 is phosphorylated *in vitro* by the p25/Cdk5 kinase at overlapping sites, but not as extensively as Pak1 phosphorylated *in vivo*. This indicates that the Pak1–Rac–p35/Cdk5 complex causes exposure and phosphorylation of additional sites on Pak1 by the p35/Cdk5 kinase. Alternatively, the presence of yet another kinase in such a complex cannot be ruled out.

To determine the effects of the p35/Cdk5 kinase on Pak1 activity, Pak1 kinase assays were done using histone H4 as an *in vitro* substrate. The Pak1 kinase induction by RacL61 was abolished in the presence of the p35/Cdk5 kinase (Fig. 4a). Titration of p35/Cdk5 revealed progressive inhibition of the activity of Pak1 kinase (Fig. 4b), whereas coexpression of Cdk5N144 with p35 had no inhibitory effects on Pak1 (Fig. 4c), indicating that active p35/Cdk5 kinase downregulates Pak1 activity. The p35/Cdk5 kinase had no effect on a constitutively active Pak1 kinase mutant (Pak1E423)

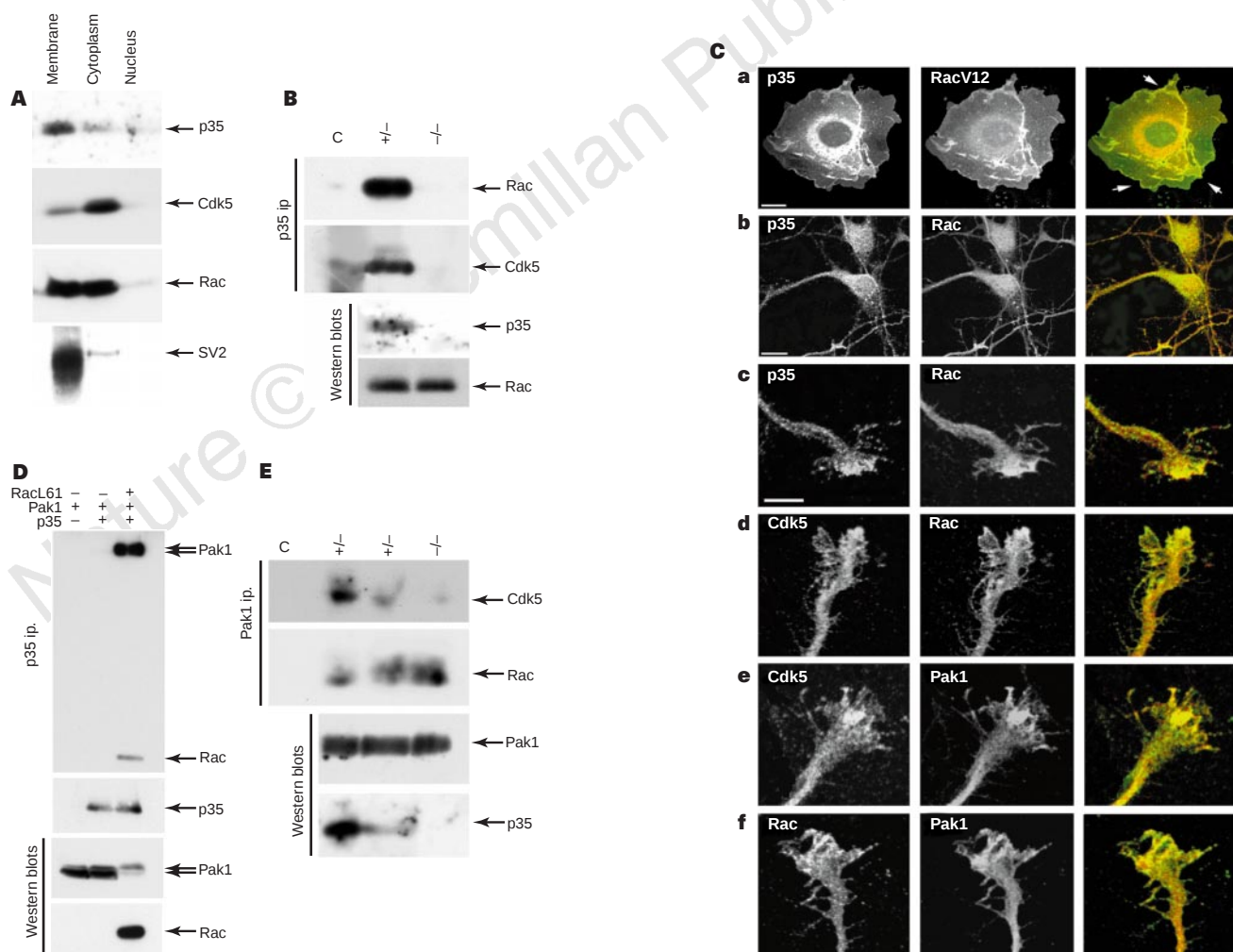


Figure 2 In neurons, p35 co-localizes and associates with Rac and Pak1 at the cell periphery. **A**, p35, Cdk5, Rac and SV2 Western blots of fractionated cortical lysates. **B**, p35 immunoprecipitates from p35^{+/+} and p35^{-/-} cortices examined for Rac or Cdk5 association. **Ca**, Swiss3T3 cells expressing p35 (red) and HA-RacV12 (green); **b**, **c**, primary neuronal cultures immunostained for p35 (green) and Rac (red); **d**, Cdk5 (red) and Rac (green); **e**, Cdk5 (green) and Pak1 (red); or **f**, Rac (green) and Pak1 (red). Arrows point to areas of p35 and Rac co-localization. Bars,

10 μ m (**a**, **b**) and 5 μ m (**c**–**f**). **D**, Lysates from COS-7 cells expressing as shown, immunoprecipitated with anti-p35 antibodies and immunoblotted with the 9E10 antibody. **E**, As in **B**, but Pak1 immunoprecipitations were subjected to Cdk5 and Rac Western blotting; including p35^{+/+} samples. In **B** and **E**, **C** depicts control immunoprecipitations; in **B**, **D**, **E**, lower panels show protein levels in individual lysates.

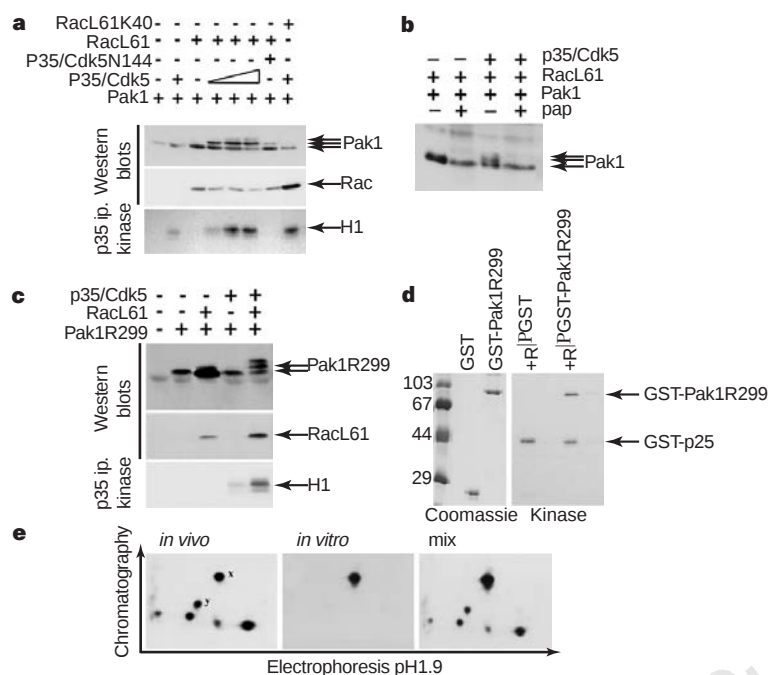


Figure 3 The p35/Cdk5 kinase associates with and phosphorylates Pak1 in a Rac-dependent manner. **a**, Western blots showing the effects of the active or inactive p35/Cdk5 kinase on myc-Pak1 in transfected COS-7 cells. **b**, Pak1 immunoprecipitates from a sample treated (+) or not (-) with potato acid phosphatase (PAP). **c**, As in **a** using the inactive Pak1R299 mutant. A small shift of unknown origin was seen in the presence of RacL61. Active Rac upregulates p35/Cdk5 activity by unknown mechanisms (**a**, **c**). **d**, GST-Pak1R299 or GST were used as substrates of the p25/Cdk5 kinase obtained from baculovirus infected Hi5 cells. The kinase was inhibited with 0.2 μ M roscovitine (R). **e**, Comparison of two-dimensional tryptic-phosphopeptide maps obtained from GST-Pak1R299 phosphorylated in COS-7 cells in the presence of p35/Cdk5 and RacL61 (*in vivo*) or phosphorylated *in vitro* by the p25/Cdk5 kinase. X and Y depict co-migrating peptides.

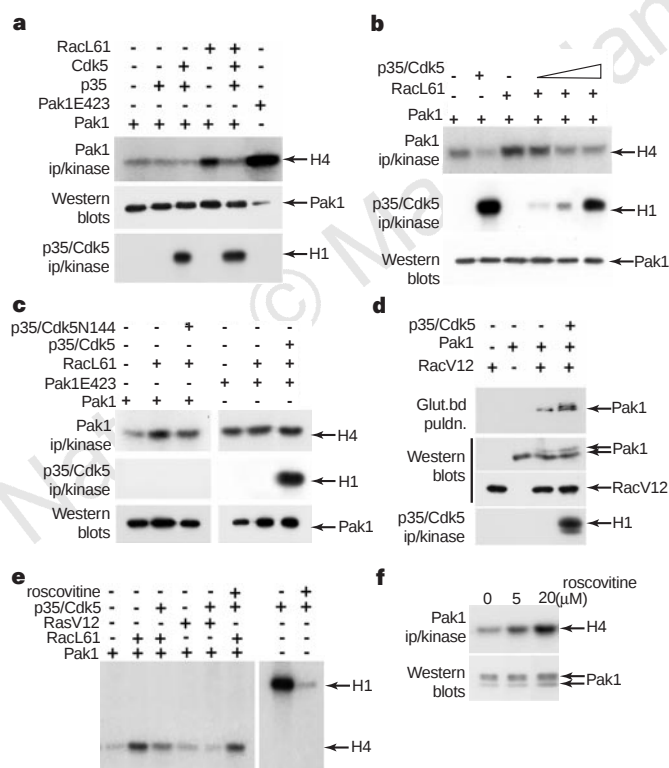


Figure 4 The p35/Cdk5 kinase inhibits Pak1 activity but not its affinity for Rac. **a-c**, Membrane lysates from COS-7 cells expressing as shown were immunoprecipitated with antibody 9E10 to isolate Pak1, and were then subjected to histone H4 kinase assays. **d**, Lysates from COS-7 cells expressing as shown were subjected to glutathione bead pulldown to isolate Rac and examined for Pak1 association by 9E10 western blots. **e**, Set amounts of Pak1 and p35 immunoprecipitates were mixed and examined for Pak1 kinase activity in the presence or absence of RacL61 or RasV12 and the inhibitor roscovitine. **f**, Lysates from primary neurons treated with roscovitine were examined for Pak1 kinase activity. Protein levels were examined by western blotting (**a-d** and **f**) and p35/Cdk5 kinase activity by phosphorylation of histone H1 (**a-e**).

(Fig. 4c). To determine whether the observed decrease in the activity of Pak1 kinase could be a consequence of reduced Rac and Pak1 association after phosphorylation by the p35/Cdk5 kinase, the extent of Pak1 and RacL61 association was compared in the presence or absence of the p35/Cdk5 kinase. Figure 4d shows that Pak1 association with Rac is unaltered, if not increased, when the p35/Cdk5 kinase is present. We further tested if Pak1 kinase activity could be inhibited by the p35/Cdk5 kinase *in vitro*. Purified Pak1 was mixed with recombinant RacL61 and the p35/Cdk5 kinase. Figure 4e demonstrates the stimulation of Pak1 kinase by RacL61, although the stimulated Pak1 activity was reduced by the p35/Cdk5 kinase. This inhibition was alleviated by roscovitine, confirming that the effect was dependent on the catalytic activity of the p35/Cdk5 kinase. To determine whether the endogenous Pak1 kinase is inhibited by p35/Cdk5 in neurons, primary cortical cultures were treated with increasing amounts of roscovitine for 1 h. A progressive rise in Pak1 kinase activity was observed, with increasing amounts of roscovitine (~fourfold at 20 μ M concentrations; Fig. 4f). Together, these data indicate that the active p35/Cdk5 kinase serves to inhibit the Pak1 kinase in neurons.

In summary, the interaction of the p35/Cdk5 kinase, Rac and Pak1 and the subsequent modification of Pak1 demonstrate the existence of a neuron-specific signalling complex. The function of the Rho GTPases and Pak1 in the regulation of the actin cytoskeleton and cell shape is now well established^{16-18,27,28}. Furthermore, it has been suggested that Pak1-mediated actin reorganization in fibroblasts has two components, only one of which requires kinase activity¹⁶. We propose that the p35/Cdk5-Pak1-Rac complex serves to inhibit, both temporally and spatially, the activity of Pak1 kinase in neurons, in this way switching Pak1 between an active or inactive state and thus allowing two distinct Pak1 activities on the actin cytoskeleton. The dynamics of this process are provided by the instability of the protein interactions, as they rely on the presence of active Rac that rapidly exchanges GTP for GDP and p35 has a short half-life²⁹. Our experiments indicate that the steady-state association between p35/Cdk5, Rac and Pak1 is small, which is due largely to the transient nature of the interactions. We therefore propose that this dynamic protein assembly functions in the rapid regulation of the actin cytoskeleton, which is necessary for neurite outgrowth and the migration of developing neurons. □

Methods

Transient expression in cells. COS-7 cells were transfected using the calcium phosphate precipitation method and Swiss3T3 using lipofectamine (GIBCO BRL). Cells were collected 24 h after transfection. The experiments used GST-, haemagglutinin- or myc-tagged GTPases, myc or GST-tagged Pak1, p35 and Cdk5 expression constructs.

Immunocytochemistry. Cells were fixed in 4% paraformaldehyde for 10 min at room temperature, followed by permeabilization with 0.3% triton X-100. Immunostaining was done as described previously⁸ using anti-p35 (pAb neu-cyc, pAb N20; Santa Cruz), anti-Cdk5 (mAb DC39, pAb 50; Kinetek Biology Corporation), anti-Pak1 (pAb, N20, Santa Cruz), anti-HA (mAb 12CA5) and anti-Rac (mAb, Upstate Biotechnology) antibodies to detect p35, Cdk5, Pak1 and Rac. This was followed by the use of Texas-red-conjugated anti-rabbit and fluorescein isothiocyanate (FITC)-conjugated anti-mouse antibodies in all cases except those in which p35 immunostaining was performed. In these cases, biotin-conjugated anti-rabbit and Texas-red-conjugated anti-mouse antibodies were used followed by FITC-conjugated streptavidin to amplify the signals. Coverslips were mounted in ProLong antifade (Molecular Probes) and analysed using a Leica or Zeiss confocal microscope.

Protein interactions. For p35/Cdk5–Rac or Pak1–Rac associations, transfected cells were lysed in 20 mM Tris pH 7.2, 2 mM MgCl₂, 0.5% NP-40, 150 mM NaCl, 1 mM DTT, protease and phosphatase inhibitors (buffer A). Lysates were incubated with glutathione beads or with 9E10 antibody and protein G sepharose or anti-p35 antibody and protein A sepharose, at 4°C, washed three times in lysis buffer and resuspended in 1 × sample buffer. For Pak1–Rac–p35/Cdk5, tissues were lysed in STM buffer (10 mM Tris HCl pH 8.0, 0.25 M sucrose, 10 mM MgCl₂, 1 mM DTT, protease and phosphatase inhibitors) containing 0.5% NP-40. Immunoprecipitations were made using as a control normal rabbit serum, anti-p35 (pAb neu-cyc, pAb C19; Santa Cruz), or anti-Pak1 (pAb, N20; Santa Cruz) followed by protein A sepharose. The beads were washed in buffer A and resuspended in 1 × sample buffer. *In vitro* interactions were done by incubating 40 µg each of GST–p35, GST–Cdk5 and GST or 10 µg GST–Pak1R299 with glutathione beads. After washing, the samples were halved and each half was incubated with equal amounts of [³²P]-GTP-labelled recombinant RacL61 or RacL61K40, from which the GST tags had been removed by thrombin cleavage, as described²⁰, at 4°C for 1 h. The beads were washed extensively in buffer A and subjected to Cerenkov counting.

Fractionation experiments. Rat embryo cortices were lysed in STM buffer. Extracts were spun at 600g and then at 100,000g to obtain cytoplasmic fractions. Membranes were extracted from the 600g spin by STM lysis with 0.5% NP-40. Nuclear fractions were the final 600g pellet lysed in RIPA (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 0.5% DOC, 0.1% SDS, 1 mM DTT, protease and phosphatase inhibitors) buffer.

Phosphorylation of Pak1. COS-7 cells were transfected with myc–Pak1 and, where appropriate, RacL61. The titration points of the p35/Cdk5 kinase are 0.5 µg, 1 µg and the maximum 2 µg each of p35 and Cdk5. For phosphatase treatment, Pak1 immunoprecipitates were resuspended in 60 µl 0.1 M MES pH 6.0, halved and one half treated with potato acid phosphatase (Sigma) for 15 min at 37°C. For *in vitro* phosphorylation of Pak1, GST–Pak1R299 was obtained from transfected COS-7 cells by glutathione-bead pull-down, eluted and dialysed against PBS. GST–p25 and GST–Cdk5 were obtained from baculovirus-infected Hi5 insect cells, immobilized on glutathione beads, and subjected to kinase assays as described previously⁸, using 1 µg of GST–Pak1R299 as a substrate. The assays were done in the presence or absence of 0.2 µM roscovitine. Two-dimensional tryptic-phosphopeptide mapping was performed comparing GST–Pak1R299 isolated from transfected COS-7 cells coexpressing p35, Cdk5 and RacL61, which had been labelled for 3 h with ³²P-orthophosphate, and GST–Pak1R299 phosphorylated *in vitro* by the GST–p25/GST–Cdk5 kinase as described previously³⁰. Equal counts of the *in vivo* or *in vitro* samples (determined by Cerenkov counting) were analysed alone or in combination to examine the extent of co-migration between individual peptides.

Kinase assays. RacL61 (10 ng) was co-transfected into COS-7 cells with Pak1 (2 µg) or Pak1E423 (2 µg) and, where appropriate, p35 (2 µg) and Cdk5 (2 µg) or Cdk5N144 (2 µg) expression vectors. In the p35/Cdk5 titration experiments, 0.1 µg, 1 µg and 3 µg each of p35 and Cdk5 were used. For Pak1 kinase assays,

membrane fractions were used. p35/Cdk5 kinase assays from lysates of transfected cells were done as described previously⁸. Pak1 kinase assays were done in the presence of 1 mM ATP, 1–2 µCi [³²P]-ATP, 3 µg histone H4 and kinase buffer (50 mM HEPES, pH 7.5, 10 mM MgCl₂) for 10 min at 30°C. For *in vitro* Pak1 kinase assays, standardized amounts of Pak1 immunoprecipitates were incubated, where appropriate, with standardized amounts of p35/Cdk5 immunoprecipitates and/or 2–5 µg of bacterially produced GTPases at 4°C. Kinase assays were done in the presence or absence of 0.2 µM roscovitine. For *in vivo* roscovitine treatment, primary neuronal cultures were exposed for 1 h to 5 µM or 20 µM roscovitine, or to DMSO as a control. Endogenous Pak1 was immunoprecipitated from the lysates using a specific antibody and subjected to kinase assays as described.

Received 17 April; accepted 2 July 1998.

- Lew, J. et al. Neuronal cdc-2-like kinase is a complex of cyclin-dependent kinase 5 and a novel brain-specific regulatory subunit. *Nature* **371**, 423–425 (1994).
- Tsai, L.-H., Delalle, I., Caviness, V. S. Jr, Chae, T. & Harlow, E. p35 is a neural-specific regulatory subunit of cyclin-dependent kinase 5. *Nature* **371**, 419–423 (1994).
- Delalle, I., Bhidé, P. G., Caviness, V. S. Jr & Tsai, L.-H. Temporal and spatial patterns of expression of p35, a regulatory subunit of cyclin-dependent kinase 5, in the nervous system of the mouse. *J. Neurocytol.* **26**, 283–296 (1997).
- Tomizawa, K. et al. Localization and developmental changes in the neuron-specific cyclin-dependent kinase 5 activator (p35^{nc5a}) in the rat brain. *Neuroscience* **74**, 519–529 (1996).
- Chae, T. et al. Mice lacking p35, a neuronal specific activator of Cdk5, display cortical lamination defects, seizures, and adult lethality. *Neuron* **18**, 29–42 (1997).
- Ohshima, T. et al. Targeted disruption of the cyclin-dependent kinase 5 gene results in abnormal corticogenesis, neuronal pathology and perinatal death. *Proc. Natl Acad. Sci. USA* **93**, 11173–11178 (1996).
- Kwon, Y. T. K. & Tsai, L.-H. A novel disruption of cortical development in p35^{−/−} mice distinct from reeler. *J. Comp. Neurol.* **395**, 510–522 (1998).
- Nikolic, M., Dudek, H., Kwon, Y. T., Ramos, Y. F. & Tsai, L.-H. The cdk5/p35 kinase is essential for neurite outgrowth during neuronal differentiation. *Genes Dev.* **10**, 816–825 (1996).
- Luo, L., Liao, Y. J., Jan, L. Y. & Jan, Y. N. Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev.* **8**, 1787–1802 (1994).
- Luo, L. et al. Differential effects of the Rac GTPase on Purkinje cell axons and dendritic trunks and spines. *Nature* **379**, 837–840 (1996).
- Kozma, R., Sarnar, S., Ahmed, S. & Lim, L. Rho family GTPases and neuronal growth cone remodeling: relationship between increased complexity induced by Cdc42Hs, Rac1, and acetylcholine and collapse induced by RhoA and lysophosphatidic acid. *Mol. Cell. Biol.* **17**, 1201–1211 (1997).
- Threadgill, R., Bobb, K. & Ghosh, A. Regulation of dendritic growth and remodeling by Rho, Rac, and Cdc42. *Neuron* **19**, 625–634 (1997).
- Zipkin, I. D., Kindt, R. M. & Kenyon, C. J. Role of a new Rho family member in cell migration and axon guidance in *C. elegans*. *Cell* **90**, 883–894 (1997).
- Lim, L., Manser, E., Leung, T. & Hall, C. Regulation of phosphorylation pathways by p21 GTPases. The p21 Ras-related Rho subfamily and its role in phosphorylation signaling pathways. *Eur. J. Biochem.* **242**, 171–185 (1996).
- Sells, M. A. & Chernoff, J. Emerging from the Pak: the p21-activated protein kinase family. *Trends Cell Biol.* **7**, 162–166 (1997).
- Sells, M. A., Knaus, U. G., Bagrodia, S., Ambrose, D. M., Bokoch, G. M. & Chernoff, J. Human p21-activated kinase (Pak1) regulates actin organization in mammalian cells. *Curr. Biol.* **7**, 202–210 (1997).
- Hall, A. Rho GTPases and the actin cytoskeleton. *Science* **279**, 509–514 (1998).
- Manser, E. et al. Expression and constitutively active α-Pak reveals effects of the kinase on actin and focal complexes. *Mol. Cell. Biol.* **17**, 1129–1143 (1997).
- van den Heuvel, S. & Harlow, E. Distinct roles for cyclin-dependent kinases in cell cycle control. *Science* **262**, 2050–2054 (1993).
- Lamarche, N. et al. Rac and Cdc42 induce actin polymerization and G1 cell cycle progression independently of p65PAK and the JNK/SAPK MAP kinase cascade. *Cell* **87**, 519–529 (1996).
- Westwick, J. K. et al. Rac regulation of transformation, gene expression, and actin organization by multiple, PAK-independent pathways. *Mol. Cell. Biol.* **17**, 1324–1335 (1997).
- Kinsella, B. T., Erdman, R. A. & Maltese, W. A. Carboxyl-terminal isoprenylation of ras-related GTP-binding proteins encoded by rac1, rac2 and ra1A. *J. Biol. Chem.* **266**, 9786–9794 (1991).
- Manser, E. et al. Molecular cloning of a new member of the p21-Cdc42/Rac-activated kinase (PAK) family. *J. Biol. Chem.* **270**, 25070–25078 (1995).
- Manser, E., Leung, T., Salihuddin, H., Zhao, Z. & Lim, L. A brain serine/threonine protein kinase activated by cdc42 and rac1. *Nature* **367**, 40–46 (1994).
- Lu, W., Katz, S., Gupta, R. & Mayer, B. J. Activation of Pak by membrane localization mediated by an SH3 domain from the adaptor protein Nck. *Curr. Biol.* **7**, 85–94 (1996).
- Meijer, L. et al. Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. *Eur. J. Biochem.* **243**, 527–536 (1997).
- Dharmawardane, S., Sanders, L. C., Martin, S. S., Daniels, R. H. & Bokoch, G. M. Localization of p21-activated kinase 1 (PAK1) to pinocytotic vesicles and cortical actin structures in stimulated cells. *J. Cell Biol.* **138**, 1265–1278 (1997).
- Daniels, R. H., Hall, P. S. & Bokoch, G. M. Membrane targeting of p21-activated kinase 1 (PAK1) induces neurite outgrowth from PC12 cells. *EMBO J.* **17**, 754–764 (1998).
- Patrick, G. N., Zhou, P., Kwon, Y. T., Howley, P. M. & Tsai, L.-H. p35, the neuronal specific activator of cdk5, is degraded by the ubiquitin-proteasome pathway. *J. Biol. Chem.* (in press).
- Dynlacht, B. D., Moberg, K., Lees, J. A., Harlow, E. & Zhu, L. Specific regulation of E2F family members by cyclin-dependent kinases. *Mol. Cell. Biol.* **17**, 3867–3875 (1997).

Acknowledgements. We thank K. Moberg and J. Lees for advice and technical help with the two-dimensional gel phosphopeptide analysis; A. Hall and J. Blenis for reagents and discussions; K. Buckley for anti-SV2 antibodies; J. Settleman for Ras expression constructs and advice; L. B. Chen and N. Perrimon for the use of confocal microscopes; and Y. T. Kwon, S. Humbert, G. Patrick, L. Zukeberg and V. Tannoch for help with the manuscript. This work was partly supported by NSF and NIH grants to L.-H.T. M.N. is supported by the Medical Foundation Charles King Trust. L.-H.T. is an assistant investigator of the Howard Hughes Medical Institute, a Rita Allen Foundation scholar and a recipient of an Esther A. and Joseph Klingenstein fund.

Correspondence and requests for materials should be addressed to L.-H.T. (e-mail: lhtsai@warren.med.harvard.edu).

Interactions controlling the assembly of nuclear-receptor heterodimers and co-activators

Stefan Westin^{*†}, Riki Kurokawa^{*†}, Robert T. Nolte[‡], G. Bruce Wisely[‡], Eileen M. McInerney[§], David W. Rose[‡], Michael V. Milburn[‡], Michael G. Rosenfeld^{†§} & Christopher K. Glass^{*†}

^{*} Division of Cellular and Molecular Medicine, [†] Division of Endocrinology and Metabolism, and [§] Howard Hughes Medical Institute, Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA

[‡] Department of Structural Chemistry, GlaxoWellcome Inc., Research Triangle Park, North Carolina 27709, USA

Retinoic-acid receptor- α (RAR- α) and peroxisome proliferator-activated receptor- γ (PPAR- γ) are members of the nuclear-receptor superfamily that bind to DNA as heterodimers with retinoid-X receptors (RXRs)^{1,2}. PPAR-RXR heterodimers can be activated by PPAR or RXR ligands³, whereas RAR-RXR heterodimers are selectively activated by RAR ligands only, because of allosteric inhibition of the binding of ligands to RXR by RAR^{4,5}. However, RXR ligands can potentiate the transcriptional effects of RAR ligands in cells⁶. Transcriptional activation by nuclear receptors requires a carboxy-terminal helical region, termed activation function-2 (AF-2) (refs 7–9), that forms part of the ligand-binding pocket and undergoes a conformational change required for the recruitment of co-activator proteins, including NCoA-1/SRC-1 (refs 10–17). Here we show that allosteric inhibition of RXR results from a rotation of the RXR AF-2 helix that places it in contact with the RAR coactivator-binding site. Recruitment of an LXXLL motif of SRC-1 to RAR in response to ligand displaces the RXR AF-2 domain, allowing RXR ligands to bind and promote the binding of a second LXXLL motif from the same SRC-1 molecule. These results may partly explain the different responses of nuclear-receptor heterodimers to RXR-specific ligands.

SRC-1 and CBP are components of a co-activator complex that mediates transcriptional activation by RAR-RXR heterodimers^{13–19}. SRC-1 interacts with nuclear receptors through a central region that contains three conserved helical motifs of the consensus sequence LXXLL^{17,20–24}, and with CBP through a distinct C-terminal domain^{14,17,23–25} (Fig. 1a). In addition, nuclear receptors interact in a ligand-dependent manner with the amino terminus of CBP^{14,16}. A region of the nuclear-receptor-interaction domain of SRC-1, encompassing LXXLL helical domains (HDs) 1 and 2 (SRC-1^{633–715}), strongly inhibited the binding of CBP to RAR (Fig. 1b), indicating competition for a common binding site. When assessed on DNA-bound RAR-RXR heterodimers, effective recruitment of CBP required the presence of both the nuclear-receptor-interaction domain and the CBP-interaction domain of SRC-1 (Fig. 1c; data not shown). Consistent with this, mutation of the CBP-interaction domain of SRC-1 abolished its ability to rescue retinoic-acid-receptor function in cells microinjected with an anti-NCoA-1/SRC-1 antibody (Fig. 1d). Thus, complex assembly and transcriptional activation by RAR-RXR heterodimers requires the concerted actions of the nuclear-receptor-interaction and CBP-interaction domains of SRC-1.

We studied combinatorial effects of RAR- and RXR-specific ligands on SRC-1 interactions with RAR-RXR heterodimers by using the RAR-specific ligand TTNPB ((E)-4-[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl] benzoic acid) and the RXR-specific ligand LG268. A saturating concentration of

TTNPB (1 μ M) resulted in a strong interaction of SRC-1 with RAR-RXR heterodimers (Fig. 1e). In contrast, LG268 did not induce SRC-1 binding, consistent with allosteric inhibition of RXR by RAR^{4,5}. However, at a concentration of TTNPB (10 nM) that resulted in minimal co-activator interaction when added alone, addition of LG268 significantly potentiated the binding of full-length SRC-1 to the heterodimer (Fig. 1e). Similar results were obtained using the nuclear-receptor-interaction domain of SRC-1 (SRC-1^{633–783}) (Fig. 2a). In contrast, LG268 alone markedly induced the binding of SRC-1^{633–783} to PPAR- γ -RXR heterodimers (Fig. 2b). LG268 also acted synergistically with the PPAR- γ -specific ligand BRL49653 to induce interactions with SRC-1^{633–783} (Fig. 2b).

The ability of LG268 to enhance the binding of SRC-1 to RAR-RXR heterodimers in the presence of TTNPB indicated that the interaction of SRC-1 with RAR might relieve the allosteric inhibition of RXR. Consistent with this, binding of the RXR-specific ligand [³H]LGD1069 to RAR-RXR heterodimers was induced by the combination of TTNPB with SRC-1^{621–1207} or the corresponding

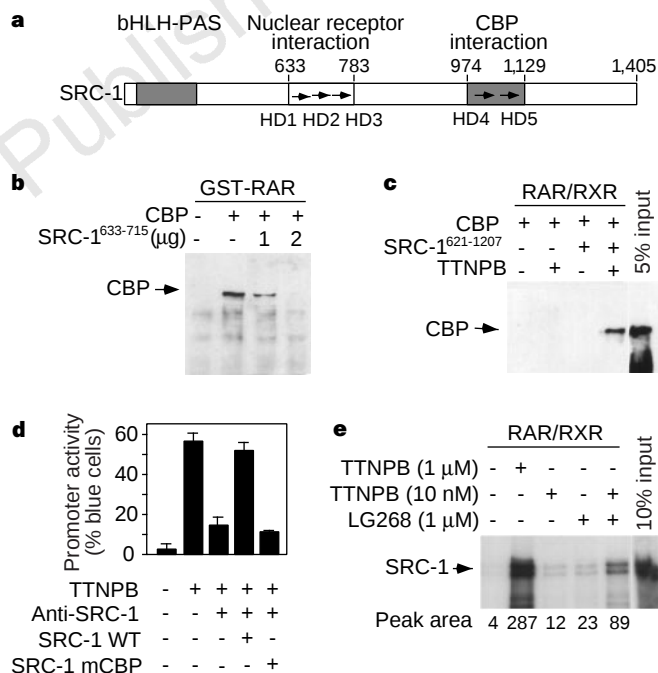


Figure 1 Recruitment of CBP to RAR-RXR heterodimers by SRC-1. **a**, Domains of SRC-1 that interact with nuclear receptors and CBP. Arrows represent helical domains (HDs) of consensus sequence LXXLL. **b**, CBP and SRC-1 compete for a common binding site on RAR. GST-RAR was incubated with full-length CBP and the RAR-specific ligand TTNPB, in the presence or absence of an SRC-1 protein fragment containing HD1 and HD2 (SRC-1^{633–715}). Specifically bound CBP was detected by western blotting. **c**, CBP is recruited to RAR-RXR heterodimers by SRC-1. Full-length CBP was incubated with RAR-RXR heterodimers bound to a biotinylated retinoic-acid-response element (RARE) in the presence of TTNPB and/or recombinant SRC-1 containing both the nuclear-receptor- and the CBP-interaction domains. Specifically bound CBP was detected as in **b**. **d**, The CBP-interaction domain is required for SRC-1 to function as a co-activator of RAR. Rat-1 cells were injected with a LacZ reporter gene under transcriptional control of an RAR-dependent promoter and treated with TTNPB as indicated. Co-injection of anti-SRC-1 IgG abolished RAR-dependent transcription, which was restored by co-injection of a plasmid directing expression of wild-type SRC-1 (SRC-1 WT), but not by expression of SRC-1 containing a mutation in the CBP-interaction domain (SRC-1 mCBP). **e**, The RXR-specific ligand LG268 potentiates the recruitment of SRC-1 to RAR-RXR heterodimers in the presence of subsaturating concentrations of the RAR-specific ligand TTNPB. RAR-RXR heterodimers were assembled on a biotinylated RARE and incubated with full-length ³⁵S-labelled SRC-1 in the presence of the indicated ligands.

region of the related co-activator P/CIP (Fig. 2c). Overexpression of SRC-1 in P19 cells markedly increased the transcriptional effects of the RXR-specific ligand LG268 in the presence of subsaturating concentrations of TTNPB, consistent with a role of SRC-1 in relief of allosteric inhibition of RXR *in vivo* (Fig. 2d).

The cooperative effects of two ligands on binding of SRC-1 to a heterodimer indicated that tandem LXXLL motifs might interact with each of the ligand-binding domains of a heterodimer. Consistent with this, SRC-1⁶³³⁻⁷¹⁵, containing HD1 and HD2, was cooperatively recruited to RAR-RXR heterodimers by LG268 and TTNPB (Fig. 3a) and to PPAR γ -RXR heterodimers by LG268 and BRL49653 (Fig. 3b). Similar results were obtained for SRC-1⁶⁵²⁻⁷⁸³, which contains HD2 and HD3 (data not shown). Mutation of either one of the two remaining helical domains in SRC-1⁶³³⁻⁷¹⁵ abolished the synergistic effects of two ligands (Fig. 3a, b), indicating a requirement for two LXXLL motifs of a single SRC-1 molecule for cooperative effects of two ligands. In the case of PPAR-RXR heterodimers, both HD1 and HD2 were required for effective interactions in the presence of a single ligand (Fig. 3b) and mutation of HD1 abolished the ability of recombinant SRC-1 to rescue PPAR function in cells microinjected with anti-SRC-1 immunoglobulin G (Fig. 3c).

As expected, both AF-2 domains of the RAR-RXR heterodimer were required so that RAR- and RXR-specific ligands could cooperatively recruit SRC-1 (Fig. 3d). Intriguingly, removal of the RAR AF-2 helix allowed some binding of SRC-1⁶³³⁻⁷¹⁵ in the presence of LG268 alone, indicating partial relief of allosteric inhibition of RXR (Fig. 3d). Deletion of the AF-2 domain from RXR resulted in a small but reproducible increase in the interaction of SRC-1⁶³³⁻⁷¹⁵ with the RAR (Fig. 3d), indicating an inhibitory role of the RXR AF-2 helix.

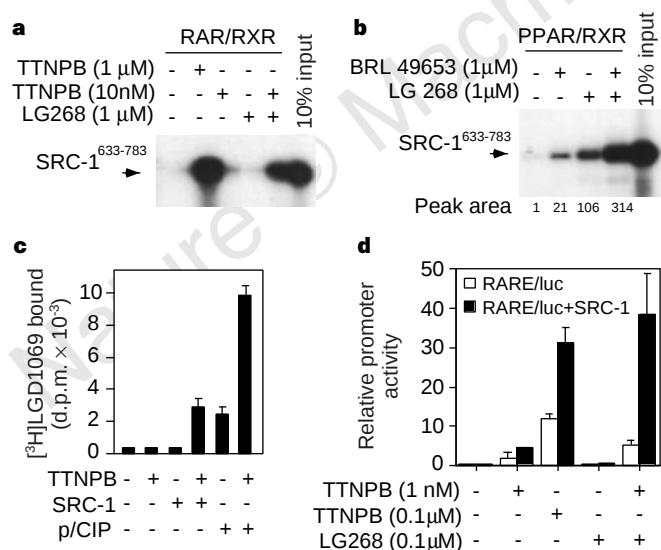


Figure 2 The nuclear-receptor-interaction domain of NCoA-1/SRC-1 allosterically regulates RXR heterodimers. **a**, Interaction of ³²P-labelled SRC-1⁶³³⁻⁷⁸³ with RAR-RXR heterodimers in the presence of a subsaturating concentration of TTNPB is potentiated by the RXR-specific ligand LG268. **b**, BRL49653 and LG268 act independently and cooperatively to recruit ³²P-labelled SRC-1⁶³³⁻⁷⁸³ to PPAR-RXR heterodimers bound to a PPAR-response element. **c**, SRC-1 and P/CIP relieve allosteric inhibition of the binding of ligands to RXR by RAR. DNA-bound RAR-RXR heterodimers were incubated with [³H]LGD1069 in the presence or absence of TTNPB and/or SRC-1⁶²¹⁻¹²⁰⁷ or P/CIP⁵⁹¹⁻¹³⁰³. **d**, NCoA-1/SRC-1 potentiates the transactivating effects of subsaturating concentrations of RAR-specific ligands in cells. P19 cells were transfected with a retinoic-acid-responsive promoter directing luciferase expression (RARE/Luc) and an expression vector for SRC-1 or a control vector as indicated, treated with the indicated concentrations of TTNPB and LG268, and analysed for luciferase activity 24 h later.

A possible explanation of these results was suggested by the crystal structure of the apo-PPAR γ ligand-binding domain: the AF-2 domain of one PPAR γ molecule interacted with the LXXLL-binding pocket of a PPAR γ molecule in a different, crystallographically related dimer²⁶. These results indicate that allosteric interactions between RAR and RXR might be explained by an interaction of the RXR AF-2 domain with the coactivator-binding site of RAR. Molecular modelling of the RAR-RXR heterodimer showed that the AF-2 motif at the end of helix 11 of RXR could easily interact with the LXXLL-binding pocket of RAR by rotation around the relatively disordered loop connecting helix 10 to helix 11 in RXR (Fig. 4a). Such an interaction would prevent closure of the ligand-binding pocket of RXR, providing a structural basis for allosteric inhibition by RAR.

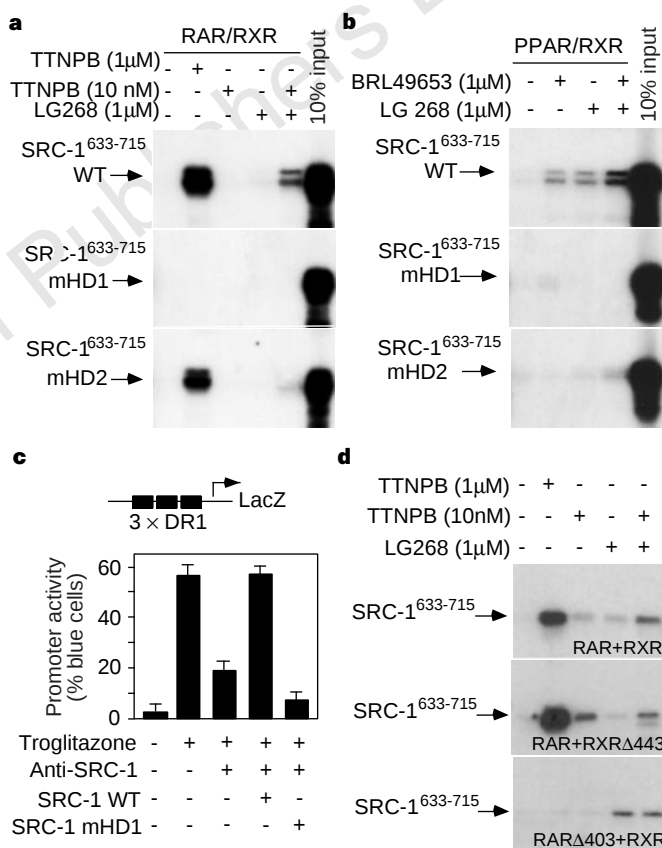


Figure 3 Structural requirements for combinatorial effects of RAR, PPAR γ and RXR ligands on interactions of NCoA-1/SRC-1 with nuclear receptors. **a**, Two helical motifs of the SRC-1 nuclear-receptor-interaction domain are required for combinatorial effects of RAR- and RXR-specific ligands. DNA-bound RAR-RXR heterodimers were incubated with ³²P-labelled SRC-1⁶³³⁻⁷¹⁵ containing HD1 and HD2 (WT), ³²P-labelled SRC-1⁶³³⁻⁷¹⁵ containing a mutation in HD1 (mHD1) or SRC-1⁶³³⁻⁷¹⁵ containing a mutation in HD2 (mHD2). **b**, Two helical motifs of the SRC-1 nuclear-receptor-interaction domain are required for combinatorial effects of PPAR γ and RXR ligands. DNA-bound PPAR γ -RXR heterodimers were incubated with the same wild-type and mutant SRC-1 proteins as in **a**. **c**, HD1 is required for SRC-1 to function as a co-activator of PPAR γ . Rat-1 cells were injected with a LacZ reporter gene under transcriptional control of a PPAR-dependent promoter and treated with troglitazone as indicated. Coinjection of anti-SRC-1 IgG abolished PPAR γ -dependent transcription, which was restored by co-injection of a plasmid directing expression of wild-type SRC-1 (SRC-1 WT), but not by a plasmid directing expression of SRC-1 containing a mutation in HD1 (SRC-1 mHD1). **d**, The AF-2 domains of both RAR and RXR are required for RXR-specific ligands to potentiate the effects of subsaturating concentrations of RAR-specific ligands on the binding of SRC-1. RAR Δ 403 and RXR Δ 443 each lack the AF-2 domain.

Other results support this model. First, DNA-bound heterodimers of RXR Δ 443 and wild-type RAR exhibited a markedly reduced ability to bind [3 H]LGD1069 in the presence of TTNPB and SRC-1 $^{633-715}$ (Fig. 4b), consistent with a requirement for the RXR AF-2 domain to close the ligand-binding pocket. Second, peptides corresponding to SRC-1 HD1 and HD2 relieved the inhibitory effects of RAR on the binding of ligands to RXR in the presence of TTNPB (Fig. 4c). Third, a peptide corresponding to the RXR AF-2 domain bound to the unliganded RAR with a higher affinity than SRC-1 HD2 (Fig. 5a), and this binding depended on the RAR AF-2 domain (data not shown). On addition of TTNPB and SRC-1 $^{633-715}$, binding of the RXR AF-2 peptide was inhibited markedly, consistent with it being displaced from the coactivator-binding site by SRC-1 (Fig. 5a). Finally, a glutathione S-transferase (GST) fusion protein containing the RXR AF-2 domain also interacted with 35 S-labelled RAR, but SRC-1 $^{633-715}$ in the presence of TTNPB effectively competed with GST-RXR-AF2 for binding to 35 S-labelled RAR (Fig.

5b). In contrast, the GST-RXR-AF2 fusion protein interacted very poorly with 35 S-labelled PPAR (Fig. 5b). Thus, the affinity of the RXR AF-2 domain for RAR and PPAR correlated with the abilities of these receptors to allosterically inhibit the binding of ligands to RXR.

These results suggest the following model for allosteric inhibition and co-activator assembly on RAR-RXR heterodimers (Fig. 5c). First, in the absence of ligand, the AF-2 domain of RXR is docked to the RAR coactivator-interaction site, preventing the binding of RXR ligands. Second, recruitment of SRC-1 through one of three LXXLL motifs in response to an RAR-specific ligand displaces the RXR AF-2 domain from RAR. Third, the release of the RXR AF-2 domain relieves allosteric inhibition, allowing ligands to bind to RXR. Fourth, the binding of an RXR ligand can then promote the interaction of a second LXXLL motif from the same SRC-1 molecule with RXR, stabilizing the complex. This model explains the selective responsiveness of RAR-RXR heterodimers to RAR ligands and

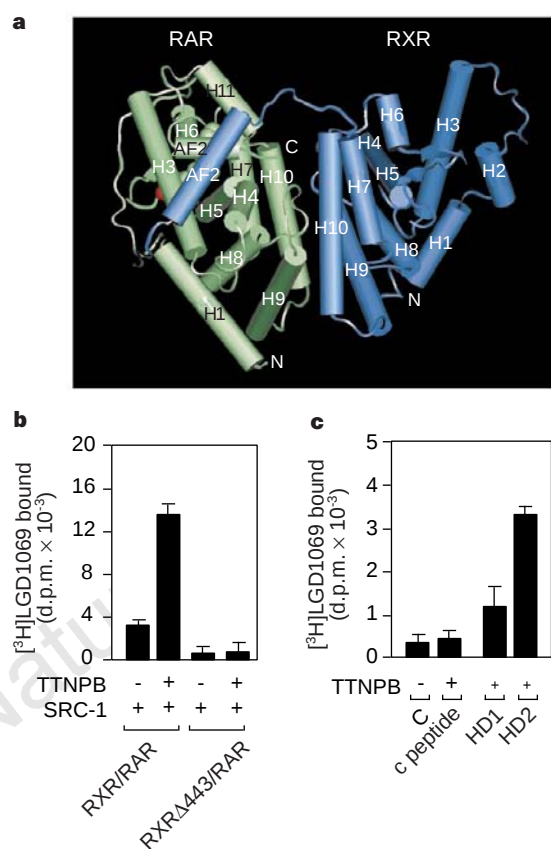


Figure 4 Interaction of an LXXLL motif with RAR relieves allosteric inhibition of RXR. **a**, Model of a heterodimer of the RAR and RXR ligand-binding domains, showing how helix 11 of RXR, containing the AF-2 domain, can dock into the activation surface of the RAR ligand-binding domain. The coordinates of RXR- α (ref. 11), shown in blue, and RAR- γ (ref. 12), shown in green, were superimposed onto the two LBD domains in the PPAR γ -SRC1 structure²⁶. The position of the RXR AF-2 helix was achieved by a rotation at residue 433, followed by a minor translation which places it against the RAR LBD in manner identical to that seen for the co-activator helix in the PPAR- γ tertiary complex. **b**, The RXR AF-2 domain is required for the binding of [3 H]LGD1069 to DNA-bound RAR-RXR heterodimers. DNA-bound RAR-RXR or RAR-RXR Δ 443 heterodimers were incubated with TTNPB as indicated, in the presence of SRC-1 $^{633-715}$. **c**, SRC-1 HD1 and HD2 peptides relieve allosteric inhibition of the binding of ligands to RXR. DNA-bound RAR-RXR heterodimers were incubated with [3 H]LGD1069 in the presence of TTNPB and the indicated SRC-1 peptides. C, no peptide added; C peptide, unrelated peptide added.



Figure 5 Mechanism of allosteric inhibition and co-activator assembly. **a**, A peptide corresponding to the AF-2 domain of RXR binds to RAR and is competed for by SRC-1 $^{633-715}$. GST-RAR was incubated with 125 I-labelled RXR AF-2 or 125 I-labelled SRC-1 HD2 peptides in the presence or absence of TTNPB and or SRC-1 $^{633-715}$. **b**, The AF-2 domain of RXR differentially interacts with RAR and PPAR. The AF-2 domain of RXR was expressed as a GST-fusion protein and tested for interaction with 35 S-labelled RAR and 35 S-labelled PPAR produced by *in vitro* translation. Binding assays were performed in the presence of activating ligands (TTNPB and BRL49653 for RAR and PPAR, respectively) and SRC-1 $^{633-715}$ as indicated (top). Control assays were performed using GST-SRC-1 $^{633-715}$ and GST alone (bottom). **c**, Proposed mechanism for relief of allosteric inhibition and co-activator assembly (see text).

resolves the apparent paradox of why RXR ligands potentiate the effects of RAR-specific ligands. The fact that RAR and PPAR show different affinities for the RXR AF-2 domain explains why RAR allosterically inhibits RXR, but PPAR does not. It is likely that this mechanism of allosteric regulation is involved in establishing receptor-specific responses of other RXR-containing heterodimers. The finding that two LXXLL motifs within a single SRC-1 molecule are used for cooperative binding to a homodimeric or heterodimeric nuclear receptor is supported by the crystal structure of SRC-1^{623–710} in a complex with a liganded PPAR- γ dimer²⁶ and is likely to be prototypic for the assembly of other complexes of nuclear receptors and co-activators. □

Methods

Assays of GST- and DNA-dependent protein-protein interactions. GST-RAR α (20–462), GST-RXR α (45–462), GST-PPAR γ (1–475) and GST-RXR-AF2(443–462) proteins were prepared as crude bacterial lysates and used for protein-protein interaction assays as described⁴. For interaction experiments with individual HD motifs and the RXR AF-2 domain, peptides were synthesized to contain a tyrosine residue at the C terminus and were labelled with ¹²⁵Iodine using Iodo-gen²⁷. 3×10^5 c.p.m. of each peptide were incubated with the liganded receptors. The sequences of the HD peptides were as follows: HD1 (amino acids 632–649), SQTSHKLVQLTTTAEQY; HD2 (amino acids 689–706), TERHKILHRLQEGSPSDY. The sequence of the RXR AF-2 (amino acids 443–462) peptide was DTPIDFTLMEMLEAPHQMTY. DNA-dependent protein-protein interaction assays using RXR heterodimers bound to biotinylated DNA response elements were performed as described⁴. Receptor-DNA complexes were incubated with 1×10^5 c.p.m. ³⁵S-labelled SRC-1 protein generated by *in vitro* transcription and translation, with 1×10^5 c.p.m. ³²P-labelled SRC-1 proteins labelled by protein kinase A, or with epitope-tagged CBP. CBP protein, containing an in-frame C-terminal seven-amino-acid extension (EYKEEEK; 'Flag' epitope), was expressed in SF9 cells in baculovirus vectors¹⁹. GST fusions of SRC-1 containing mutations in a single HD motif (LXXLL to LAAAA) were generated by polymerase chain reaction from the corresponding mutations in the pCMX-NCoA-1/SRC-1 expression vector¹⁷.

Assays of DNA-dependent ligand binding. Ligand-binding assays were performed nearly as for DNA-dependent protein-protein interaction⁴, except that we used 5 nM radiolabelled [³H]LGD1069 (6,7-[³H]LGD1069) as an RXR-specific ligand. P/CIP and NCoA-1/SRC-1 were used as GST-fusion proteins expressed in *Escherichia coli*, or were cleaved from the GST portion using thrombin.

Transient transfection and single-cell microinjection assays. Transfection experiments in P19 cells were done using the standard calcium phosphate procedure²⁸. Cells were transfected with 1 μ g of a retinoic-acid-response element/luciferase reporter and 1 μ g of either pCMX-P/CIP or pCDNA3-SRC-1 expression plasmids; the cells were then treated with TNPB or LG268 at the indicated concentrations and assayed for luciferase activity 24 h later. Microinjection assays of co-activator function were performed as described¹⁷. Briefly, plasmids were co-injected into the nuclei of Rat-1 fibroblast cells at 100 μ g ml⁻¹, together with either preimmune IgG or anti-NCoA-1/SRC-1 IgG. Cells were stimulated with ligand 6 h after injection and assayed for injected IgG

and LacZ expression 16 h later. At least 200 injected cells were counted for each condition.

Received 20 June; accepted 3 August 1998.

- Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
- Glass, C. K. Differential recognition of target genes by nuclear receptor monomers, dimers and heterodimers. *Endocrinol. Rev.* **15**, 1503–1519 (1994).
- Kliwer, S. A., Umesono, K., Noonan, D. J., Heyman, R. A. & Evans, R. M. Convergence of 9-cis retinoic acid and peroxisome proliferator signaling pathways through heterodimer formation of their receptors. *Nature* **358**, 771–774 (1992).
- Kurokawa, R. *et al.* Regulation of retinoid signaling by receptor polarity and allosteric control of ligand binding. *Nature* **371**, 528–531 (1994).
- Forman, B. M., Umesono, K., Chen, J. & Evans, R. M. Unique response pathways are established by allosteric interactions among nuclear hormone receptors. *Cell* **81**, 541–550 (1995).
- Chen, J.-Y. *et al.* Two distinct actions of retinoid-receptor ligands. *Nature* **382**, 819–822 (1996).
- Durand, B. *et al.* Activation function 2 (AF2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF2 activity. *EMBO J.* **13**, 5370–5382 (1994).
- Danielian, P. S., White, R., Lees, J. A. & Parker, M. G. Identification of a conserved region required for hormone-dependent transcriptional activation by steroid hormone receptors. *EMBO J.* **11**, 1025–1033 (1992).
- Baretino, D., Vivanco Ruiz, M. M. & Stunnenberg, H. G. Characterization of the ligand-dependent transactivation domain of thyroid hormone receptor. *EMBO J.* **13**, 3039–3049 (1994).
- Wagner, R. L. *et al.* A structural role for hormone in the thyroid hormone receptor. *Nature* **378**, 690–697 (1995).
- Bourguet, W., Ruffin, M., Chambon, P., Gronemeyer, H. & Moras, D. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR- α . *Nature* **375**, 377–382 (1995).
- Renaud, J.-P. *et al.* Crystal structure of the RAR- γ ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681–689 (1995).
- Onate, S. A., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* **270**, 1354–1357 (1995).
- Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414 (1996).
- Hanstein, B. *et al.* p300 is a component of an estrogen receptor coactivator complex. *Proc. Natl Acad. Sci. USA* **93**, 11540–11545 (1996).
- Chakravarti, D. *et al.* Role of CBP/p300 in nuclear receptor signaling. *Nature* **383**, 99–103 (1996).
- Torchia, J. *et al.* The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387**, 677–684 (1997).
- Korzus, E. *et al.* Transcription factor-specific requirements for coactivators and their acetyltransferase functions. *Science* **279**, 703–707 (1998).
- Kurokawa, R. *et al.* Differential use of CREB binding protein-coactivator complexes. *Science* **279**, 700–703 (1998).
- Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733–736 (1997).
- Ding, X. F. *et al.* Nuclear receptor-binding sites of coactivators glucocorticoid receptor interacting protein 1 (GRIP1) and steroid receptor coactivator 1 (SRC-1): multiple motifs with different binding specificities. *Mol. Endocrinol.* **12**, 302–313 (1998).
- Le Douarin, B. *et al.* A possible involvement of TIF1 α and TIF1 β in the epigenetic control of transcription by nuclear receptors. *EMBO J.* **15**, 6701–6715 (1996).
- Voegel, J. J. *et al.* The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *EMBO J.* **17**, 507–519 (1998).
- Kalkhoven, E., Valentine, J. E., Heery, D. M. & Parker, M. G. Isoforms of steroid receptor co-activator 1 differ in their ability to potentiate transcription by the oestrogen receptor. *EMBO J.* **17**, 232–243 (1998).
- Yao, T.-P., Ku, G., Zhou, N., Scully, R. & Livingston, D. M. The nuclear hormone receptor coactivator SRC-1 is a specific target of p300. *Proc. Natl Acad. Sci. USA* **93**, 10626–10631 (1996).
- Nolte, R. T. *et al.* Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- γ . *Nature* (in the press).
- Fraker, P. J. & Speck, J. C. J. Protein and cell membrane iodinations with a sparingly soluble chloroamide, 1,3,4,6-tetrachloro-3a 6a-diphenylglycoluril. *Biochem. Biophys. Res. Commun.* **80**, 849–857 (1978).
- Chen, C. & Okayama, H. High efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752 (1987).

Acknowledgements. We thank R. Heyman for making [³H]LGD1069 available, S. Green for help with radio-iodination of peptides and T. Schneiderman for help with manuscript preparation. S.W. was supported by grants from The Swedish Cancer Society and a training grant from the NIH. M.G.R. acknowledges support from the HHMI. C.K.G. is an Established Investigator of the American Heart Association. This work was also supported by grants from the NIH (to C.K.G. and M.G.R.).

Correspondence and requests for materials should be addressed to C.K.G. (e-mail: cglass@ucsd.edu).